AD-A080 446 TECHNOLOGY SERVICE CORP SANTA MONICA CALIF F/G 12/1

NEW METHODS FOR ESTIMATING TAIL PROBABILITIES AND EXTREME VALUE--ETC(U)

DEC 79 L BREIMAN, C J STONE, J D GINS F49620-79-C-0171

UNCLASSIFIED TSC-PD-A226-1 AFOSR-TR-80-0019 NL

1 OF 2

AD
A080446

AFOSR-TR- 80-0019

3 LEVEL II

AD A080446

DDC FILE COPY

DDC
RECEIVED
FEB 4 1980
B

80 1 29 138

Technology Service Corporation

Approved for public release;
distribution unlimited.

Technology Service Corporation

2811 WILSHIRE BOULEVARD • SANTA MONICA, CALIFORNIA 90403 • PH. (213) 829-7411

3 LEVEL II

18 AFOSR

6

NEW METHODS FOR ESTIMATING
TAIL PROBABILITIES AND
EXTREME VALUE DISTRIBUTIONS.

19 TR-80-0019

14 TSC-PD-A226-1

11 14 December 1979

12 103

9 Final rept.,

10 Leo
Charles
John

L./Breiman
C. J./Stone
J. D./Gins

FINAL REPORT
Contract F49620-79-C-0171

15

16 2304 17 A5

Submitted to:

Dr. Ismail Shimi
U.S. Air Force
Office of Scientific Research
Bolling Air Force Base
Washington, D.C. 20332

404132

CONTENTS

1. INTRODUCTION

Let X denote a positive random variable whose underlying distribution function F is continuous and strictly increasing on $[0,\infty]$. For $x \geq 0$ let $S(x)$ be the tail (survival) probability defined by $S(x) = P(X>x) = 1-F(x)$. For $0<p<1$ let x_p denote the (upper) p^{th} quantile of X defined by

$$S(x_p) = 1-F(x_p) = P(X>x_p) = p .$$

In this report methods are studied for estimating tail probabilities and quantiles based on a random sample of size n from the distribution of X. Interest will center around problems of estimating $S(x)$ and x_p when n is large and $S(x)$ and p are both quite small. Good solutions to these estimation problems are particularly useful in connection with extreme value statistics. For the probability that the maximum of a random sample of size k from the distribution of X exceeds x is equal to $1 - (1-S(x))^k$; and the q^{th} quantile of this sample maximum is equal to x_p, where $p = 1 - (1-q)^{1/k}$. Observe that if k is large, then p is quite small - in fact, approximately equal to $k^{-1} \log(1/(1-q))$.

Global parametric estimators typically do not provide good solutions to these estimation problems. For they are nonrubust to departures from the assumed model in the tails of the distribution of X, which departures are present in most applications. Standard nonparametric procedures, based directly on empirical tail probabilities suffer from lack of sufficient data beyond the point x_p of interest especially if, say, $p=1/n$. In this

1. INTRODUCTION

Let X denote a positive random variable whose underlying distribution function F is continuous and strictly increasing on $[0,\infty]$. For $x \geq 0$ let $S(x)$ be the tail (survival) probability defined by $S(x) = P(X>x) = 1-F(x)$. For $0<p<1$ let x_p denote the (upper) p^{th} quantile of X defined by

$$S(x_p) = 1-F(x_p) = P(X>x_p) = p .$$

In this report methods are studied for estimating tail probabilities and quantiles based on a random sample of size n from the distribution of X. Interest will center around problems of estimating $S(x)$ and x_p when n is large and $S(x)$ and p are both quite small. Good solutions to these estimation problems are particularly useful in connection with extreme value statistics. For the probability that the maximum of a random sample of size k from the distribution of X exceeds x is equal to $1 - (1-S(x))^k$; and the q^{th} quantile of this sample maximum is equal to x_p, where $p = 1 - (1-q)^{1/k}$. Observe that if k is large, then p is quite small - in fact, approximately equal to $k^{-1} \log(1/(1-q))$.

Global parametric estimators typically do not provide good solutions to these estimation problems. For they are nonrubust to departures from the assumed model in the tails of the distribution of X, which departures are present in most applications. Standard nonparametric procedures, based directly on empirical tail probabilities suffer from lack of sufficient data beyond the point x_p of interest especially if, say, $p=1/n$. In this

report estimators will be proposed and studied which are a compromise between completely parametric and completely nonparametric procedures.

The estimation problems discussed here pertain to the right tail of the distribution of X. But the same methods are applicable to the left tail, which is also important in applications. In particular if X is a positive random variable, its left tail can be estimated by applying the procedures described here to the right tail of 1/X. Our interest in problems of estimating tail probabilities and extreme value distributions arose initially in environmental applications (see Breiman, Stone and Gins [3]), where they are connected with violation of standards. But the same problems also occur in reliability, strength of materials, fatique and many other applications. (See Section 6.3 of Gumbel [6] and Section 3.12 of Galambos [5]).

The specific research objectives of this work are as follows:

a. Define exponential tail and transformed exponential tail procedures for estimating probabilities in the upper tail of an underlying distribution and for estimating the corresponding quantiles.

b. Study analytically the bias and variance of the estimators and, for the transformed exponential tail procedure, verify the existence and uniqueness of the estimators.

c. Design and run a Monte Carlo experiment to supplement the analytic study of the bias and variance of the various estimators for a variety of sample sizes, underlying distributions and parameters in the procedures.

d. Based on the analytic and Monte Carlo results draw appropriate conclusions and make suggestions for further research.

2. GENERAL DISCUSSION

In Section 2.1 the main estimators considered in this report are defined and some of their basic properties are summarized. In Section 2.2 some alternative estimators are noted. A Monte Carlo experiment has been designed to compare all these estimators under a variety of sample sizes and underlying distributions. The results of this experiment are presented in Section 2.3. In Section 2.4 conclusions are drawn based on the research done to date and suggestions are made for further work. More detailed discussions of the material in this section are given in Sections 3 through 6.

2.1 EXPONENTIAL TAIL AND WEIBULL TAIL ESTIMATORS

Let the values obtained in a random sample of size n from the distribution of X be written in decreasing order as $X_1 > \ldots > X_n$.

Suppose first that X has an exponential distribution with a mean a. Then the tail probabilities are given by $S(x) = e^{-x/a}$, $x \geq 0$. The maximum likelihood estimator of the parameter a is given by $\hat{a} = (X_1 + \ldots + X_n)/n$. This leads to the estimator $\hat{S}(x) = e^{-x/\hat{a}}$ of S(x).

Let x_0 be a positive number. If X is exponential with mean a, then the conditional distribution of $X - x_0$ given that $X > x_0$ is exponential with mean a. Suppose the latter but not necessarily the former property holds. Then $S(x) = S(x_0)e^{-(x-x_0)/a}$, $x \geq x_0$. The parameter $S(x_0)$ can be estimated by $\hat{S}(x_0) = n^{-1}\ \#\{i : X_i \geq x_0\}$, where # I denotes the number of elements in the set I. Set $m = \#\{i : X_i > x_0\}$. If $m \geq 1$, then maximum

likelihood considerations suggest estimating a by $\hat{a} = m^{-1}\sum_1^m (X_i - x_0)$. This now leads to the estimator $\hat{S}(x)$ of $S(x)$ given by

$$\hat{S}(x) = \hat{S}(x_0)e^{-(x-x_0)/\hat{a}} \quad , \quad x \geq x_0 \quad .$$

In practice a positive integer m is chosen such that $2<m<n$ and the preceeding estimator is applied with $x_0 = X_{m+1}$. That is,

$$\hat{a} = \frac{1}{m}\sum_1^m (X_i - X_{m+1}) \tag{2.1}$$

and

$$\hat{S}(x) = \left(\frac{m+1}{n}\right) e^{-(x-X_{m+1})/\hat{a}} \quad , \quad x \geq X_{m+1} \quad . \tag{2.2}$$

Associated with (2.2) is the estimate of x_p given by

$$\hat{x}_p = X_{m+1} + \hat{a}\log\frac{m+1}{np} \quad , \quad 0<p \leq \frac{m+1}{n} \quad . \tag{2.3}$$

In order to complete the definition of $\hat{S}(x)$ let the empirical estimator be be used for $x<X_{m+1}$; that is, set

$$\hat{S}(x) = \frac{1}{n}\#\{i:X_i \geq x\} \quad , \quad 0 \leq x < X_{m+1} \quad . \tag{2.4}$$

Similarly, to complete the definition of $\hat{x}_p$ set

$$\hat{x}_p = X_{\lceil(np)}, \quad \frac{m+1}{n} < p \leq 1 \quad , \tag{2.5}$$

where $\lceil(y)$ denotes the smallest integer which equals or exceeds y (a notation borrowed from APL). The estimators given by (2.2) and (2.3) are

called exponential tail estimators. They have been used successfully in a number of applications, starting with Breiman, Stone and Gins [3]. Very similar estimators have been proposed by Hill [7] and Weissman [18].

Hopefully when (2.2) and (2.3) are applicable they should yield more accurate estimates than the empirical estimators given by (2.4) and (2.5). Suppose in particular that $p=1/n$, let $\hat{x}_{1/n}$ be the estimator of the $(1/n)^{th}$ quantile $x_{1/n}$ given by (2.3) and let $M_n = X_1$ be the estimator of $x_{1/n}$ given by (2.5). (Note that M_n is the maximum value in the random sample of size n.) It is reasonable to expect that for some choice of m as a function of n with $m\to\infty$ and $m/n\to 0$ as $n\to\infty$

$$\lim_{n\to\infty} \frac{E(\hat{x}_{1/n} - x_{1/n})^2}{E(M_n - x_{1/n})^2} = 0 \quad .$$

Theorem 6.3 contains a more general result for a large class of underlying distribution functions F.

Suppose now that X^b has an exponential distribution with mean a for some $b>0$ (so that X itself has a Weibull distribution) and let x_0 be a positive number. Then the conditional distribution of $X^b - x_0^b$ given that $X>x_0$ is exponential with mean a. Thus

$$P(X>x|X>x_0) = P(X^b - x_0^b > x^b - x_0^b|X>x_0) = e^{-(x^b-x_0^b)/a} \quad , \; x\geq x_0 \; ,$$

and hence

$$S(x) = S(x_0)e^{-(x^b-x_0^b)/a} \quad , \quad x\geq x_0 \quad . \qquad (2.6)$$

This suggests estimators of S(x) of the form given by

$$\hat{S}(x) = \hat{S}(x_0)e^{-(x^{\hat{b}}-x_0^{\hat{b}})/\hat{a}} \quad , \quad x \geq x_0 \quad .$$

As before, replace x_o by X_{m+1} and set $\hat{S}(x_0) = (m+1)/n$. This leads to the estimator of $S(x)$ given by

$$\hat{S}(x) = \frac{m+1}{n} \; e^{-(x^{\hat{b}}-X_{m+1}^{\hat{b}})/\hat{a}} \quad , \quad x \geq X_{m+1} \quad . \qquad (2.7)$$

Associated with (2.7) is the estimator of x_p given by

$$\hat{x}_p = \left(X_{m+1}^{\hat{b}} + \hat{a} \log \frac{m+1}{np}\right)^{1/\hat{b}} \quad , \quad 0<p\leq \frac{m+1}{n} \quad . \qquad (2.8)$$

The estimators given by (2.7) and (2.8) are called Weibull tail estimators. They are supplemented by (2.4) and (2.5) respectively. (For similar estimators see Weinstein [16] and followup papers by Jeruchim [8] and Filimonov [4].)

Of course in order to complete the definitions of Weibull tail estimators, it is necessary to specify $\hat{a}$ and $\hat{b}$. Given b, maximum likelihood considerations suggest estimating a by

$$\hat{a} = \frac{1}{m}\sum_{1}^{m} \left(X_i^{\hat{b}} - X_{m+1}^{\hat{b}}\right) \quad .$$

There are a variety of reasonable ways to choose $\hat{b}$. One approach is maximum likelihood based on an assumed Weibull tail form for the underlying distribution F. But this approach leads to numerical solution of an equation with no particularly nice properties (such as always having a unique solution). Also, the maximum likelihood approach is not theoretically compelling unless the assumed theoretical form is really believed to be valid.

Another approach is to make some feature of the empirical distribution of $(X_i^{\hat{b}} - X_{m+1}^{\hat{b}})/\hat{a}$, $1 \leq i \leq m$, match up with the corresponding feature of the standard exponential distribution. The test for exponentiality due to Shapiro and Wilk [14] (see also Stephens [15]) suggests making the empirical second moment of $(X_i^{\hat{b}} - X_{m+1}^{\hat{b}})/\hat{a}$, $1 \leq i \leq m$, equal to the second moment of the standard exponential distribution, namely 2. According to this criterion, $\hat{b}$ should be chosen to satisfy the equation.

$$U(\hat{b}) = 2 \quad , \tag{2.9}$$

where

$$U(b) = \frac{\frac{1}{m}\sum_1^m \left(X_i^b - X_{m+1}^b\right)^2}{\left[\frac{1}{m}\sum_1^m \left(X_i^b - X_{m+1}^b\right)\right]^2} \quad .$$

It is shown in Section 4.2 that U is a continuous strictly increaasing function on $(0,\infty)$ which has limits at 0 and ∞ given by $U(\infty) = m > 2$ and

$$U(0) = \frac{\frac{1}{m}\sum_1^m \log^2 \frac{X_i}{X_{m+1}}}{\left(\frac{1}{m}\sum_1^m \log \frac{X_i}{X_{m+1}}\right)^2} \quad .$$

Thus (2.9) has at most one solution. It does have a solution if and only if

$$\frac{\frac{1}{m}\sum_{1}^{m} \log^2 \frac{X_i}{X_{m+1}}}{\left(\frac{1}{m}\sum_{1}^{m} \log \frac{X_i}{X_{m+1}}\right)^2} < 2 \quad . \tag{2.10}$$

Since U is continuous and strictly increasing--indeed, it has a strictly positive derivative--on $(0,\infty)$ it is easy to solve (2.9) numerically when (2.10) holds. When (2.10) fails to hold the estimators

$$\hat{S}(x) = \left(\frac{m+1}{n}\right)\left(\frac{X_{m+1}}{x}\right)^{1/\hat{a}} \quad , \qquad x \geq x_0, \tag{2.11}$$

and

$$\hat{x}_p = X_{m+1}\left(\frac{m+1}{np}\right)^{\hat{a}} \quad , \qquad 0 < p \leq S(x_0), \tag{2.12}$$

with

$$\hat{a} = \frac{1}{m}\sum_{1}^{m} \log \frac{X_i}{X_{m+1}}$$

are used instead of the Weibull tail estimators. The estimators given by (2.11) and (2.12) are called <u>Pareto tail</u> estimators (see Hill [7]). They arise as limits of Weibull tail estimators by letting $\hat{b} \to 0$.

A general form of transformed exponential tail estimators is discussed in Section 4.1. This form includes exponential tail, Pareto tail and Weibull tail estimators as special cases. Other special cases are mentioned in Section 4.3. One such special case, which provides a generalization of the Pareto tail estimators, is always applicable when (2.10) fails to hold. (For another fairly general approach to the estimation of tail probabilities, see Pickands [12].)

In a modified setting in Section 5 an optimality property is shown to hold for Weibull tail estimators when the underlying distribution F is a tail Weibull distribution; i.e., when (2.6) holds for some x_0 and some choice of the positive parameters a and b.

2.2 OTHER ESTIMATORS

Given $1 \leq N < 100$, exponential tail and Weibull tail estimators with $m = \lceil(Nn/100)$ are referred to as exponential tail N% and Weibull tail N% estimators. In the Monte Carlo experiment these estimators were compared with various other estimators, which will now be described.

The empirical estimators of $S(x)$ and x_p are defined by

$$\hat{S}(x) = \frac{1}{n} \#\{i : X_i \geq x\} \quad , \qquad x \geq 0, \tag{2.13}$$

and

$$\hat{x}_p = X_{\lceil(np)} \quad , \qquad 0 < p \leq 1. \tag{2.14}$$

If X^b has an exponential distribution with mean a (so that X itself has a Weibull distribution), then $S(x) = e^{-x^b/a}$. This suggests estimators $\hat{S}(x)$ of the form

$$\hat{S}(x) = e^{-x^{\hat{b}}/\hat{a}} \quad , \qquad x \geq 0 \quad , \tag{2.15}$$

where the positive numbers $\hat{a}$ and $\hat{b}$ are chosen from the sample data. The corresponding estimator of x_p is given by

$$\hat{x}_p = (\hat{a} \log \frac{1}{p})^{1/\hat{b}} \quad , \quad 0<p\leq 1 \quad . \tag{2.16}$$

The parameters $\hat{a}$ and $\hat{b}$ can be chosen by the maximum likelihood method. According to this method

$$\hat{a} = \frac{1}{n}\sum_1^n X_i^{\hat{b}} \quad , \tag{2.17}$$

while $\hat{b}$ satisfies the equation

$$\frac{\sum_1^n X_i^{\hat{b}} \log X_i}{\sum_1^n X_i^{\hat{b}}} - \frac{1}{\hat{b}} = \frac{1}{n} \sum_1^n \log X_i \quad . \tag{2.18}$$

The right side of (2.18) is continuous and strictly increasing in $\hat{b}$. (indeed,the derivative is strictly positive) on $(0,\infty)$ and has limit $-\infty$ as $\hat{b}\to 0$ and limit $\log X_1 > n^{-1}\sum_1^n \log X_i$ as $\hat{b}\to\infty$. Thus (2.18) has a unique solution, which is easily found numerically. The estimators given by (2.15) and (2.16) with $\hat{a}$ and $\hat{b}$ given by (2.17) and (2.18) are called Weibull maximum likelihood estimators.

For an alternative method of choosing $\hat{a}$ and $\hat{b}$, pick m and m' with $1\leq m'<m<n$. Choose $\hat{a}$ and $\hat{b}$ such that $\hat{S}(X_{m'}) = m'/n$ and $\hat{S}(X_m) = m/n$. These two equations can be solved explicitly, yielding the formulas

$$\hat{a} = \frac{X_m^{\hat{b}}}{\log \frac{n}{m}} \tag{2.19}$$

and

$$\hat{b} = \frac{\log \log \frac{n}{m'} - \log \log \frac{n}{m}}{\log X_{m'} - \log X_m} . \tag{2.20}$$

The estimators given by (2.15) and (2.16) with $\hat{a}$ and $\hat{b}$ given by (2.19) and (2.20) are called Weibull percentile estimators. (Precisely, (2.15) and (2.16) are applied when $x \geq X_m$ and $0<p \leq m/n$ respectively; otherwise, the empirical estimators given by (2.13) and (2.14) are used.) For $1 \leq N < 100$ the Weibull percentile estimators with $m = \lceil(Nn/100)$ and $m' = \lceil(Nn/1000)$ are referred to as Weibull N% estimators.

If log X has a normal distribution with mean μ and variance σ^2 (so that X itself has a lognormal distribution), then $S(x) = Q(\sigma^{-1}(\log x-\mu))$, where

$$Q(x) = \frac{1}{2\pi} \int_x^\infty e^{-y^2/2} dy .$$

This suggest estimators of S(x) of the form

$$\hat{S}(x) = Q\left(\frac{\log x-\hat{\mu}}{\hat{\sigma}}\right) , \quad x>0 , \tag{2.21}$$

where $\hat{\mu}$ and $\hat{\sigma}>0$ are chosen from the sample data. The corresponding estimator of x_p is given by

$$\hat{x}_p = e^{\hat{u}+\hat{\sigma}\, Q^{-1}(p)} \quad , \qquad 0<p<1. \tag{2.22}$$

The numbers $\hat{\mu}$ and $\hat{\sigma}$ can be chosen by the maximum likelihood method, which yields

$$\hat{\mu} = \frac{1}{n}\sum_{1}^{n} \log X_i \tag{2.23}$$

and

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{1}^{n} (\log X_i - \hat{\mu})^2 \; . \tag{2.24}$$

The estimators given by (2.21) and (2.22) with $\hat{\mu}$ and $\hat{\sigma}$ given by (2.23) and (2.24) are called <u>lognormal maximum likelihood</u> estimators.

For an alternative method of choosing $\hat{\mu}$ and $\hat{\sigma}$, pick m and m' with $1 \le m' < m < n$. Choose $\hat{\mu}$ and $\hat{\sigma}$ such that $\hat{S}(X_{m'}) = m'/n$ and $\hat{S}(X_m) = m/n$. These two equations can be solved explicitly, which yields the formulas

$$\hat{\mu} = \log X_m - \hat{\sigma}\, Q^{-1}\left(\frac{m}{n}\right) \tag{2.25}$$

and

$$\hat{\sigma} = \frac{\log X_{m'} - \log X_m}{Q^{-1}\left(\frac{m'}{n}\right) - Q^{-1}\left(\frac{m}{n}\right)} \quad . \tag{2.26}$$

The estimators given by (2.21) and (2.22) with $\hat{\mu}$ and $\hat{\sigma}$ given by (2.25) and (2.26) are called <u>lognormal percentile</u> estimators. (Precisely, (2.21)

and (2.22) are applied when $x \geq X_m$ and $0<p\leq m/n$ respectively; otherwise the empirical estimators given by (2.13) and (2.14) are used.) For $1\leq N<100$ lognormal percentile estimators with $m= \lceil(Nn/100)$ and $m'= \lceil(Nn/1000)$ are referred to as lognormal N% estimators.

Let a pair of estimators $\hat{S}(x)$ and $\hat{x}_p$ be written as $\hat{S}(x; X_1 \ldots, X_n)$ and $\hat{x}_p(X_1,\ldots,X_n)$ to denote explicitly the dependence on the sample data. The pair is said to be <u>scale invariant</u> if for all choices of $x>0$, $0<p<1$, $c>0$ and $X_1>\ldots>X_n>0$

$$\hat{S}(cx, cX_1,\ldots cX_n) = \hat{S}(x,X_1,\ldots,X_n)$$

and

$$\hat{x}_p(cX_1,\ldots,cX_n) = c\hat{x}_p(X_1,\ldots,X_n).$$

The pair is said to be <u>power invariant</u> if for all choices of $x>0$, $0<p<1$, $c>0$ and $X_1>\ldots>X_n>0$

$$\hat{S}(x^c,X_1^c,\ldots,X_n^c) = \hat{S}(x,X_1,\ldots,X_n)$$

and

$$\hat{x}_p(X_1^c,\ldots,X_n^c) = (\hat{x}_p(X_1,\ldots,X_n))^c .$$

All of the above estimators are scale invariant, and all except exponential tail estimators are power invariant.

2.3 MONTE CARLO COMPARISONS

A Monte Carlo experiment was performed to determine the accuracy of exponential tail and Weibull tail estimators of tail probabilities and quantiles and to compare these estimators with the other estimators described in Section 2.2. Specifically the following twelve estimators were evaluated:

empirical;

exponential tail 5%, 10%, 25%;

Weibull tail 25%, 50%;

Weibull maximum likelihood;

Weibull percentile 10%, 25%;

lognormal maximum likelihood;

lognormal percentile 10%, 25%.

Twenty underlying distribution functions were considered. They will be defined in terms of four groups each having five distribution functions, the distribution functions in the i^{th} group being denoted by F_{ij}, $1 \leq j \leq 5$. The distributions functions in each group will depend on five values b_j, $1 \leq j \leq 5$, of a parameter b given explicitly by $b_1=.5$, $b_2=.75$, $b_3=1$, $b_4=1.5$ and $b_5=2$.

The first group consists of Weibull distribution functions, defined by

$$F_{ij}(x) = 1 - e^{-x^{b_j}}, \quad x \geq 0 \quad .$$

In particular F_{13} is the standard exponential distirbution function, given by $F_{13}(x) = 1-e^{-x}$, $x \geq 0$. Let X be a random variable having distribution

function F_{13}. Then X^{1/b_j} has distribution function F_{1j} for $1 \leq j \leq 5$. Let $X_1 > \ldots > X_n$ be the ordered sample values based on a random sample of size n from the standard exponential distribution function F_{13}. Then $X_1^{1/b_j} > \ldots > X_n^{1/b_j}$ can be thought of as the ordered sample values based on a random sample of size n from F_{1j}. This observation, together with the power invariance of all but exponential tail estimators, leads to substantial computational savings in the Monte Carlo experiment. The same observation and savings are applicable to each of the other three groups of five distribution functions defined below.

The second group consists of mixture of the corresponding Weibull distributions in the first group. Specifically let X now be a random variable which with probability 1/2 is chosen from an exponential distribution with mean 1/3 and with probability 1/2 is chosen from an exponential distribution with mean 5/3. Then X has mean 1 as before. For $1 \leq j \leq 5$, let F_{2j} be the distribution function of X^{1/b_j}; then

$$F_{2j}(x) = \frac{1}{2}[F_{ij}(3^{b_j} x) + F_{ij}((\tfrac{3}{5})^{b_j} x)], \quad x \geq 0 \quad .$$

The third group consists of lognormal distributions. Specifically let X now be a lognormal random variable having unit mean and variance. For $1 \leq j \leq 5$ let F_{3j} be the distribution function of X^{1/b_j}. If X_j has distribution function F_{3j}, then log X_j is normally distributed with mean μ_j and variance σ_j^2 given by

$$\mu_j = - \frac{\log 2}{2b_j} \qquad \text{and} \qquad \sigma_j^2 = \frac{\log 2}{b_j^2};$$

also

$$F_{3j}(x) = \Phi\left(\frac{\log x - \mu_j}{\sigma_j}\right) \quad , \qquad x>0 \ .$$

The fourth group consists of mixtures of the corresponding lognormal distributions in the third group. Specifically for $1 \leq j \leq 5$,

$$F_{4j}(x) = \tfrac{1}{2}[F_{33}(3^{b_j}x) + F_{33}((\tfrac{3}{5})^{b_j}x] \ , \qquad x>0 \ .$$

In particular

$$F_{43}(x) = \tfrac{1}{2}[F_{33}(3x) + F_{33}(\tfrac{3x}{5})], \qquad x>0 \ .$$

If X has distribution function F_{43}, then X^{1/b_j} has distribution function F_{4j}.

Weibull and lognormal distributions were chosen because they are commonly employed in reliability, environmental and other applications where extreme values play an important role. The indicated mixtures were included to give a larger diversity of shapes. They were motivated in part by the useful role that normal mixtures have played in robustness studies of estimators of a location parameter. The selected values .5, .75, 1, 1.5 and 2 of the parameter b were chosen based on experience in fitting Weibull distributions to environmental data.

It is useful to define a measure of the heaviness of the distribution function of the random variable X in order to classify the distribution function as heavy-tailed, medium-tailed or light-tailed. To this end, suppose that the function ℓ defined by $\ell(x) = -\log P(X>x) = -\log S(x)$ is twice differentiable and has positive first derivative on $(0,\infty)$. It is convenient to require that the

measure of heaviness be local and, specifically, that it depend only on the first two derivatives of ℓ evaluated at a fixed quantile x_p, where $0<p<1$; let the measure be denoted by $H_X(p)$. It is natural to require that $H_X(p)$ be scale invariant; i.e., that $H_{cX}(p) = H_X(p)$ for $c>0$. It is also natural to require that if X is exponentially distributed, then $H_X(p)=0$, $H_{X^{1/b}}(p)>0$ for $0<b<1$, and $H_{X^{1/b}}(p)<0$ for $b>1$ (note that $X^{1/b}$ has a Weibull distribution with shape parameter b).

Define r on $(0,\infty)$ by

$$r(x) = \frac{\ell''(x)}{(\ell'(x))^2}, \qquad x>0,$$

and set

$$H_X(p) = -r(x_p).$$

Then $H_X(p)$ satisfies all the criteria mentioned in the previous paragraph. It is easily seen that

$$H_{X^{1/b}}(p) = \frac{1-b}{bx_p\ell'(x_p)} + H_X(p) \quad \text{for } b>0. \tag{2.27}$$

In particular, if X is exponentially distributed, then

$$H_{X^{1/b}}(p) = \frac{1-b}{b\log(1/p)} \qquad \text{for } b>0.$$

Observe here that $H_{X^{1/b}}(p)$ converges to zero slowly as $p\to\infty$. If $\log X$ is normally distributed with variance σ^2, so that X itself is lognormally distributed, the convergence of $H_X(p)$ to zero as $p\to 0$ is even slower; specifically, by straightforward calculations,

$$H_X(p) = (\sigma+z_p)\,\frac{p}{\phi(z_p)} - 1 \sim \frac{\sigma}{(2\log\frac{1}{p})^{\frac{1}{2}}} \qquad \text{as} \quad p\to 0,$$

where ϕ is the density of the standard normal distribution function Φ and $1 - \Phi(z_p) = p$.

The values of $H_X(p)$ corresponding to the twenty underlying distributions used in the Monte Carlo experiment are shown in Table 1. The different rows correspond to the different groups, indicated under "F" by the following abbreviations: WI (Weibull), WM (Weibull mixture), LN (lognormal mixture), and LN (lognormal). The different columns correspond to the indicated values for the parameter b. (By (2.27), within each group the heaviness is a linear function of b^{-1} with positive slope.)

According to Table 1 among the twenty underlying distribution functions there is a preponderance of heavy-tailed distribution functions. The three heaviest-tailed distribution functions F_{21}, F_{31} and F_{41} have much heavier tails than those which typically arise in the environmental applications in which exponential tail estimators were originally developed. The relative performance of exponential tail estimators in the numerical comparisons which follow would be significantly improved if results for these three distribution functions were ignored in making the comparisons.

Table 1. Heaviness of distribution functions at $x_{.1}$

F	.50	.75	1.00	1.50	2.00
WI	.43	.14	.00	-.14	-.22
WM	.64	.23	.03	-.18	-.28
LM	.80	.41	.22	.02	-.07
LN	.68	.36	.20	.05	-.03

In the Monte Carlo experiment the sample size n took on the values 100, 200, 400 and 800. The quantiles x_p were estimated for p = 1/50, 1/100, 1/200, 1/400, 1/800 and 1/1600. Tail probabilities were estimated

as x ranged over the quantiles x_p $(p=1/50,\ldots,1/600)$ of whatever underlying distribution was being simulated at the time.

Two different measures were used initially to evaluate estimators of tail probabilities, namely

$$R_{ij}^2(\hat{S}(x_p)) = E_{ij}(\hat{S}(x_p) - S(x_p))^2$$

and

$$L_{ij}^2(\hat{S}(x_p)) = E_{ij}(\log \hat{S}(x_p) - \log S(x_p))^2 \quad .$$

Here E_{ij} refers to expectation when F_{ij} is the underlying distribution function and x_p is the p^{th} quantile of F_{ij}. Note that the second measure L_{ij}^2 is inapplicable to the empirical estimator of $S(x_p)$ since the estimate has positive probability of being zero. The measure

$$R_{ij}^2(\hat{x}_p) = \frac{E_{ij}(\hat{x}_p - x_p)^2}{E_{ij}(X_{\Gamma(pn)} - x_p)^2}$$

was used initially to evaluate estimators of quantiles. Note that $R_{ij}^2(\hat{x}_p)$ is the ratio of the mean square error of the estimator $\hat{x}_p$ to the mean square error of the empirical estimator $X_{\Gamma(pn)}$ of x_p. The use of this measure was motivated in part by the asymptotic results for exponential tail estimators obtained in Section 6.

Monte Carl estimates of the above measures were obtained for each estimator and each quantity to be estimated by averaging over a number of

replications. Specifically for sample sizes 100, 200, 400 and 800 there were respectively 600, 400, 300 and 200 replications. (The relative standard errors (see Section 3.3) of $R_{ij}(\hat{S}(x_p))$, $L_{ij}(\hat{S}(x_p))$ and $R_{ij}(\hat{x}_p)$ were typically about .05 or less. This is small enough to warrant the qualitative conclusions which follow.)

Summary statistics were then obtained by averaging these measures over the twenty underlying distribution functions and then, for ease in presentation, taking square roots. The results are given in Section 3.

Next, in light of the summary statistics in Section 3.3, it was decided to look at five estimators for each sample size: empirical; exponential tail N% for the best overall value of N (for the given sample size n and p=1/n); Weibull tail 50%; Weibull percentile N% for the best overall value of N; and lognormal percentile N% for the best overall value of N. The specific values of N% are as follows:

Estimator	n			
	100	200	400	800
Exponential tail	25%	10%	10%	5%
Weibull percentile	25%	25%	10%	10%
lognormal percentile	25%	25%	10%	10%

For each sample size n and for p=1/n the quantities

$$\min_{\hat{S}(x_p)} R_{ij}(\hat{S}(x_p)),$$

$$\min_{\hat{S}(x_p)} L_{ij}(\hat{S}(x_p)),$$

and

$$\min_{\hat{x}_p} \quad R_{ij}(\hat{x}_p)$$

were evaluated. Here the first and third minimizations were over all five estimators selected for the given sample size; while the second minimization involved all of these estimators except the empirical estimator, for which $L_{ij}(\hat{S}(x_p))$ isn't applicable. Corresponding "efficiencies" were defined by

$$e_{ij}(\hat{S}(x_p)) = \frac{\min\limits_{\hat{S}(x_p)} \quad R_{ij}(\hat{S}(x_p))}{R_{ij}(\hat{S}(x_p))} \quad ,$$

$$e^L_{ij}(\hat{S}(x_p)) = \frac{\min\limits_{\hat{S}(x_p)} \quad L_{ij}(\hat{S}(x_p))}{L_{ij}(\hat{S}(x_p))} \quad ,$$

and

$$e_{ij}(\hat{x}_p) \quad = \frac{\min\limits_{\hat{x}_p} \quad R_{ij}(\hat{x}_p)}{R_{ij}(\hat{x}_p)} \quad .$$

(Note that the usual definition of efficiency is the square of the definition used here. Note also that the measure $e^L_{ij}(\hat{S}(x_p))$ is not applicable to the empirical estimator.) These efficiencies were then averaged over the twenty

underlying distribution functions, which yielded measures denoted by $\overline{e}(\hat{S}(x_p))$, $\overline{e}^{L}(\hat{S}(x_p))$ and $\overline{e}(\hat{x}_p)$. The results are shown in Table 2. Here the abbreviations EM (empirical), LP (lognormal percentile), WP (Weibull percentile), WT (Weibull tail) and ET (exponential tail) are used for the five estimators.

Relative to the average efficiency $\overline{e}(\hat{x}_p)$ of the estimators of the quantile x_p when $p = 1/n$, the exponential tail estimator does essentially as well as or better than any of the other estimators for each of the four sample sizes considered. This rather surprising result lends much credence to exponential tail estimators of $x_{1/n}$. The Weibull tail estimator comes out a close second. It should be noted that the exponential tail estimator does substantially better than the empirical estimator, but that this improvement is rather insensitive to the sample size n. The theoretical results in Section 6 suggest increasing improvement with increasing sample size.)

Relative to the average efficiency $\overline{e}(\hat{S}(x_p))$ of estimators of the tail probability S(x) when this probability actually equals 1/n, the Weibull tail estimator overall does the best, while the Weibull percentile estimator comes out a close second and the estimator tail estimator is third best but still performs reasonably well. Surprisingly different results are obtained when the average efficiency $\overline{e}^{L}(\hat{S}(x_p))$ is used instead of $\overline{e}(\hat{S}(x_p))$. Now the lognormal percentile estimator does best for each of the four sample sizes; the Weibull tail estimator is a close second for the two larger sample sizes, but there is no good competitor to the lognormal percentile estimator for the two smaller sample sizes.

Table 2. Average efficiencies for p=1/n

n	Measure	ET	WT	WP	LP	EM
100	$\bar{e}(S(\hat{x}_p))$	.88	.97	.86	.66	.71
	$\bar{e}^L(S(\hat{x}_p))$	.74	.53	.61	.94	
	$\bar{e}(\hat{x}_p)$	.97	.86	.81	.49	.41
200	$\bar{e}(S(\hat{x}_p))$	.84	.98	.92	.60	.68
	$\bar{e}^L(S(\hat{x}_p))$	.74	.76	.69	.95	
	$\bar{e}(\hat{x}_p)$	.96	.92	.89	.51	.42
400	$\bar{e}(S(\hat{x}_p))$	.87	.92	.95	.57	.63
	$\bar{e}^L(S(\hat{x}_p))$	.74	.85	.60	.88	
	$\bar{e}(\hat{x}_p)$	.94	.92	.83	.49	.41
800	$\bar{e}(S(\hat{x}_p))$	.84	.91	.91	.57	.59
	$\bar{e}^L(S(\hat{x}_p))$	.72	.86	.66	.89	
	$\bar{e}(\hat{x}_p)$	.90	.91	.80	.54	.35

It was next decided to determine how the efficiency depends on the choice of the underlying distribution function. Except for exponential tail estimators the various efficiencies e_{ij} are relatively insensitive to j and can be approximated satisfactorily by $(e_{i1} + \ldots + e_{i5})/5$. This should be expected, since for the other estimators $R_{ij}(\hat{S}(x_p))$ and (when applicable) $L_{ij}(\hat{S}(x_p))$ are independent of j and

$$R_{ij}(\hat{x}_p) = \frac{E_{i3}\left(\hat{x}_p^{1/b_j} - x_p^{1/b_j}\right)^2}{E_{i3}\left(X_{\Gamma(pn)}^{1/b_j} - x_p^{1/b_j}\right)^2} \quad .$$

For exponential tail estimators, however, e_{ij} varies considerably with j. Thus we decided to calculate $(e_{i1} +\ldots+ e_{i5})/5$ for each of the (four or) five estimators and in addition to calculate e_{ij}, $1\leq j\leq 5$, for exponential tail estimators. The results are shown in Table 3. The efficiencies e_{ij}, $1\leq j\leq 5$, for the exponential tail estimators are shown in the first five columns, the headings denoting the corresponding values of b_j. The entries under the column labeled "F" describe the distribution functions in the i^{th} group: WI (Weibull), WM (Weibull mixture), LM (lognormal mixture) and LN (lognormal).

Table 3 sheds light mainly on the behavior of exponential tail estimators. It is worthwhile to look at the first five columns of Table 3 in conjunction with Table 1. As expected, exponential tail estimators perform very well when the heaviness of the underlying distribution function is near zero.

Table 3. Individual efficiencies for p=1/n

n	Measure	F	.50	.75	1.00	1.50	2.00	ET	WT	WP	LP	EM
100	$e(\hat{S}(x_p))$	WI	.91	1.00	.99	.85	.76	.90	.99	.69	.65	.72
		WM	.83	.98	.98	.80	.68	.86	1.00	.85	.60	.68
		LM	.56	.84	.97	1.00	.95	.86	.95	1.00	.69	.72
		LN	.58	.89	1.00	1.00	1.00	.89	.96	.88	.71	.74
	$e^L(\hat{S}(x_p))$	WI	.40	.75	1.00	1.00	1.00	.83	.48	.55	.91	
		WM	.26	.55	.85	1.00	1.00	.73	.64	.73	.96	
		LM	.22	.43	.66	1.00	1.00	.66	.48	.61	.96	
		LN	.29	.54	.77	1.00	1.00	.72	.53	.54	.93	
	$e(\hat{x}_p)$	WI	1.00	1.00	1.00	.91	.77	.94	.92	.69	.51	.49
		WM	1.00	1.00	1.00	.92	.72	.93	.89	.75	.40	.51
		LM	1.00	1.00	1.00	1.00	1.00	1.00	.80	.91	.48	.34
		LN	1.00	1.00	1.00	1.00	1.00	1.00	.81	.89	.58	.32
200	$e(\hat{S}(x_p))$	WI	.84	.91	.91	.87	.84	.87	1.00	.70	.57	.67
		WM	.84	.94	.93	.85	.79	.87	.97	1.00	.48	.65
		LM	.49	.74	.85	.90	.90	.78	.93	1.00	.63	.66
		LN	.56	.80	.90	.96	.96	.84	1.00	.97	.70	.72
	$e^L(\hat{S}(x_p))$	WI	.51	.77	.94	1.00	1.00	.84	.75	.73	.96	
		WM	.39	.66	.85	1.00	1.00	.78	.98	.90	.86	
		LM	.29	.49	.67	.91	1.00	.67	.66	.57	.99	
		LN	.31	.51	.67	.88	1.00	.68	.65	.57	1.00	
	$e(\hat{x}_p)$	WI	1.00	1.00	1.00	.89	.82	.94	.97	.76	.49	.52
		WM	1.00	1.00	1.00	.94	.82	.95	.89	.88	.35	.49
		LM	.85	1.00	1.00	1.00	1.00	.97	.92	.95	.54	.32
		LN	.87	1.00	1.00	1.00	1.00	.97	.90	.98	.66	.36

Table 3. (Cont.) Individual efficiencies for p=1/n

n	Measure	F	.50	.75	1.00	1.50	2.00	ET	WT	WP	LP	EM
400	$e(\hat{S}(x_p))$	WI	.91	1.00	.94	.77	.68	.86	1.00	.85	.59	.66
		WM	.91	1.00	1.00	.77	.63	.86	.83	.97	.54	.65
		LM	.59	.82	.95	1.00	.96	.86	.93	1.00	.53	.61
		LN	.70	.86	.99	1.00	1.00	.91	.94	.97	.61	.61
	$e^L(\hat{S}(x_p))$	WI	.49	.89	1.00	1.00	1.00	.88	.85	.55	.88	
		WM	.30	.63	.92	.97	.87	.74	1.00	.74	.77	
		LM	.22	.42	.64	1.00	1.00	.65	.79	.53	.95	
		LN	.26	.48	.71	1.00	1.00	.69	.78	.57	.93	
	$e(\hat{x}_p)$	WI	1.00	1.00	1.00	.82	.69	.90	.97	.77	.51	.44
		WM	1.00	1.00	1.00	.86	.67	.91	.83	.92	.45	.52
		LM	.96	.99	1.00	1.00	1.00	.99	.94	.76	.44	.32
		LN	.87	.94	1.00	1.00	1.00	.96	.93	.85	.58	.34
800	$e(\hat{S}(x_p))$	WI	.85	.98	.93	.79	.70	.85	1.00	.84	.52	.58
		WM	.88	.98	.90	.73	.64	.83	.64	1.00	.43	.57
		LM	.60	.76	.87	.91	.86	.80	1.00	.80	.53	.53
		LN	.60	.83	.95	1.00	.99	.88	1.00	.99	.81	.67
	$e^L(\hat{S}(x_p))$	WI	.44	.78	1.00	1.00	1.00	.84	.97	.69	.85	
		WM	.38	.70	.92	1.00	.96	.79	.96	1.00	.74	
		LM	.22	.40	.60	.92	1.00	.63	.84	.47	.98	
		LN	.23	.40	.57	.83	.99	.60	.68	.48	1.00	
	$e(\hat{x}_p)$	WI	.88	1.00	.98	.80	.69	.87	1.00	.78	.48	.37
		WM	1.00	1.00	1.00	.80	.67	.89	.68	.97	.36	.40
		LM	.73	.81	.91	1.00	.96	.88	1.00	.63	.53	.31
		LN	.88	.95	1.00	1.00	1.00	.97	.96	.82	.79	.32

It is remarkable how slowly the behavior of the exponential tail estimator of the quantile $x_{1/n}$ deteriorates as the heaviness parameter becomes positive. The behavior of the exponential tail estimator of the tail probability $\hat{S}(x_{1/n})$ relative to the criterion $L_{ij}(\hat{S}(x_{1/n})$ deteriorates much more rapidly as the heaviness becomes positive; this is because the indicated criterion severely penalizes substantial underestimates of tail probabilities. Observe that much different behavior occurs if the criterion $R_{ij}(\hat{S}(x_{1/n})$, which severely penalizes substantial overestimates of tail probabilities, is used instead of $L_{ij}(\hat{S}(x_{1/n})$.

The above analysis of the first five columns of Table 3 suggests a new procedure which hopefully combines the best features of exponential tail and Weibull tail estimators. Namely, first choose the parameter $\hat{b}$ according to the Weibull tail 50% estimator, and then apply the exponential tail N% estimator to the transformed sample data $X_1^{\hat{b}},\ldots,X_n^{\hat{b}}$. This two-step procedure is power invariant as well as scale invariant, so its efficiency is insensitive to the parameter b. Thus in trying to "guestimate" its efficiency it can be assumed that b=1.

Suppose the underlying distribution is Weibull or a Weibull mixture with b=1. It should then turn out that $\hat{b}$ is approximately equal to one, in which case the two-step procedure should behave similarly to the exponential tail N% procedure. But inspection of Table 3 shows that this estimator does uniformly well for all four sample sizes and all three criteria when b=1 and the distribution is Weibull or a Weibull mixture. Suppose instead that the underlying distribution is lognormal or a lognormal mixture. It should

then turn out that $\hat{b}$ is approximately equal to .5. Thus the two-step procedure should behave similarly to the exponential tail N% procedure when b=2. Again, inspection of Table 3 shows that this estimator does uniformly well when b=2 and the distribution is lognormal or a lognormal mixture.

There is room for further improvement when using the two-step procedure. Recall that in Table 2 and Table 3 for the exponential tail N% estimator, N% takes on the values 25%, 10%, 10% and 5% respectively as n ranges over the sample sizes 100, 200, 400 and 800. If larger values of N% were chosen the decrease in variance would generally not compensate for the increase in bias. The two-step procedure, however, should be much less biased than the exponential tail procedure, so larger values of N% may yield still further improvement for the two-step procedure.

The two-step procedure looks very promising at this point in time for the reasons just described. But further speculation along these lines is unnecessary, since it will be included in the next Monte Carlo experiment, which will be started soon.

2.4 CONCLUSIONS AND DIRECTIONS FOR FURTHER RESEARCH

This report has focused on two procedures, exponential tail and Weibull tail, for estimating tail probabilities $S(x)$ and quantiles x_p based on a random sample of size n form an underlying distribution. Interest has centered around estimating $S(x)$ when $S(x) = 1/n$ and estimating x_p when $p = 1/n$.

The theoretical results that were obtained in Sections 4-6, mostly of an asymptotic nature, suggest that both procedures should be reasonably effective in estimating the quantities of interest. More precise numerical results were obtained by means of a Monte Carlo experiment. These results, summarized in Section 2.3, confirm that both procedures are promising, but strongly suggest that substantial further improvements could be made by properly combining them.

In the next phase of this research, interest will first be focused on combining the desirable features of exponential tail and Weibull tail procedures to obtain a two-step procedure which (hopefully) performs well uniformly over all the situations considered in Section 2.3. The two-step procedure should lead to nearly unbiased estimators in these situations. If so, an effort will be made to attach approximately valid confidence intervals to these estimators, such confidence intervals being very desirable in applications; for this,the theory in Section 6 should be helpful.

The two-step procedure is scale and power invariant. However, in some situations, for example, when the underlying distribution is believed to be nearly normal or when a logarithmic transformation is first applied to the sample data, it is appropriate to use procedures which are location and scale invariant. Modifications of the two-step procedure which possess these two invariance properties will be studied.

The numerical results presented in Tables 2 and 3 of Section 2.3 shown that the efficiency of estimators of tail probabilities is surprisingly sensitive to the choice of measure used in evaluating this efficiency. So

far two measures, $E(\hat{S}(x) - S(x)^2$ and $E\left(\log \frac{\hat{S}(x)}{S(x)}\right)^2$ have been employed. It is desirable also to include a measure which is directly related to the problem of estimating extreme value distributions. Recall that the distribution function for the maximum of a random sample of size n from the underlying distribution is of the form $(1-S(x))^n$, $x \geq 0$, which can be estimated by $(1-\hat{S}(x))^n$, $x \geq 0$. This leads to the measure $E((1-\hat{S}(x))^n - (1-S(x))^n)^2$, which will be included in the next Monte Carlo experiment. (The new measure should be closer to $E\left(\log \frac{\hat{S}(x)}{S(x)}\right)^2$ than to $E(\hat{S}(x)-S(x))^2$.)

Finally, efforts will be made to analyze appropriately modified versions of the best procedures when the sample data is grouped, dependent, or nonstationary.

3. MONTE CARLO EXPERIMENT

The Monte Carlo experiment and its results were discussed in Section 2.3. Further details are given in Sections 3.1 and 3.2 and some additional summary statistics are presented in Section 3.3.

3.1 GENERATION OF PSEUDO-SAMPLE DATA

As indicated in Section 3.2, for $1 \leq i \leq 4$ a pseudo-random sample $X_1 > \ldots > X_n$ from F_{i3} was generated. Then for $1 \leq j \leq 5$, $X_1^{1/bj} > \ldots > X_n^{1/bj}$ was used as a pseudo-random sample from F_{ij}.

Let $(U_1, \ldots, U_n)$ and $(V_1, \ldots, V_n)$ denote disjoint sequences of consecutive uniform pseudo-random numbers generated according to the algorithm described in Lewis, Goodman and Miller [9]. Disjoint such sequences were generated for each pair (i,n) and each replication. (Actually only the U's were generated for $i=1$ and $i=3$.)

The random sample corresponding to the standard exponential distribution function F_{13} was obtained by sorting the values $-\log(1-U_m)$, $1 \leq m \leq n$. The random sample corresponding to the exponential mixture F_{23} was obtained by sorting the values

$$\left(\frac{4}{3} I(V_m) - \frac{5}{3}\right) \log (1-U_m), \quad 1 \leq m \leq n,$$

where $I(V_m)=0$ or 1 according as $V_m < .5$ or $V_m \geq .5$.

The random sample corresponding to the lognormal distribution function F_{33} was obtained by sorting the values

$$\exp(\mu_3 + \sigma_3 \Phi^{-1}(U_m)), \quad 1 \leq m \leq n,$$

where μ_3 and σ_3 are defined in Section 2.3. The inverse Φ^{-1} of the standard normal distribution function was calculated according to (26.2.23) of Abramowitz and Stegum [1]. The random sample corresponding to the lognormal mixture F_{43} was obtained by sorting the values

$$\left(\frac{4}{3} I(V_m) - \frac{5}{3}\right) \exp(\mu_3 + \sigma_3 \Phi^{-1}(U_m)), \quad 1 \leq m \leq n.$$

3.2 STANDARD ERRORS OF MONTE CARLO ESTIMATES

For each estimator T of a quantity t, Monte Carlo estimates of the bias, standard deviation and root mean square error were obtained based on the indicated number N of replications (e.g., N=600 for n=100). The results, while too extensive to report here, certainly helped lead to the conclusions in Sections 2.3 and 2.4. Formulas for approximating the standard errors of these Monte Carlo estimates will now be described.

Let μ and σ denote respectively the mean and standard deviation of T and set $\nu_2=ET^2=\mu^2 + \sigma^2$, $\nu_4=ET^4$ and $\mu_4=E(T-\mu)^4$. Note that $Var(T^2)=\nu_4 - \nu_2^2$ and $Var((T-\mu)^2)=\mu_4 - \sigma^4$. Let $T_1,\ldots,T_N$ be a random sample of size N from the distribution of T. Then the quantities just defined can be estimated by

$$\hat{\mu} = \overline{T} = \frac{T_1 + \ldots + T_N}{N} ,$$

$$\hat{\sigma} = (\frac{1}{N}\sum_1^N (T_i - \overline{T})^2)^{1/2} ,$$

$$\hat{\nu}_2 = \frac{1}{N}\sum_1^N T_i^2 = \hat{\mu}^2 + \hat{\sigma}^2 ,$$

$$\hat{\nu}_4 = \frac{1}{N}\sum_1^N T_i^4 ,$$

and

$$\hat{\mu}_4 = \frac{1}{N}\sum_1^N (T_i - \overline{T})^2 .$$

The bias μ -t of T is estimated by

$$\hat{Bias}\ (T) = \hat{\mu} - t = \overline{T} - t .$$

The standard error of this estimate is $\sigma/\sqrt{N}$. This leads to the approximation

$$\hat{SE}\ (\hat{Bias}\ (T)) = \frac{\hat{\sigma}}{\sqrt{N}} .$$

The standard deviation σ of T is estimated by

$$\hat{SD}\ (T) = \hat{\sigma} .$$

Now

$$\sqrt{N}(\hat{\sigma}^2 - \sigma^2) = \frac{1}{\sqrt{N}} \sum_1^N ((T_i - \mu)^2 - \sigma^2) - \frac{1}{\sqrt{N}} (\sqrt{N}(\overline{T} - \mu))^2 ,$$

so by the central limit theorem the distribution of $\sqrt{N}\,(\hat{\sigma}^2 - \sigma^2)$ converges to a normal distribution with mean zero and variance $\mathrm{Var}((T - \mu)^2) = \mu_4 - \sigma^2$. Consequently (by (6a.2.1) of Rao [13]) the distribution of $\sqrt{N}\,(\hat{\sigma} - \sigma)$ converges to a normal distribution with mean zero and variance $(\mu_4 - \sigma^4)/4\sigma^2$. This leads to the approximation

$$\hat{SE}(\hat{SD}(T)) = \left(\frac{\hat{\mu}_4 - \hat{\sigma}^4}{4\hat{\sigma}^2 N}\right)^{1/2} .$$

The root mean square error $((\mu - t)^2 + \sigma^2)^{1/2}$ of T is estimated by

$$\hat{RMSE}(T) = \sqrt{\hat{\nu}_2} = \sqrt{\hat{\mu}^2 + \hat{\sigma}^2} .$$

An argument similar to that used in the previous paragraph leads to the approximation

$$\hat{SE}(\hat{RMSE}(T)) = \left(\frac{\hat{\nu}_4 - \hat{\nu}_2^2}{4\hat{\nu}_2 N}\right)^{1/2} .$$

The above formulas can be applied to $\hat{S}(x_p)$ and $\hat{x}_p$ and also to $\log \hat{S}(x_p)$ viewed as an estimator of $\log S(x_p)$

3.3 SUMMARY STATISTICS

The summary statistics described briefly in Section 2.3 are given explicitly by

$$R(\hat{S}(x_p)) = \left(\frac{1}{20}\sum_{i=1}^{4}\sum_{j=1}^{5} E_{ij}(\hat{S}(x_p) - S(x_p))^2\right)^{1/2},$$

$$L(\hat{S}(x_p)) = \left(\frac{1}{20}\sum_{i=1}^{4}\sum_{j=1}^{5} E_{ij}(\log \hat{S}(x_p) - \log S(x_p))^2\right)^{1/2},$$

and

$$R(\hat{x}_p) = \left(\frac{1}{20}\sum_{i=1}^{4}\sum_{j=1}^{5} \frac{E_{ij}(\hat{x}_p - x_p)^2}{E_{ij}(X_{\lceil (pn)} - x_p)^2}\right)^{1/2}.$$

Here E_{ij} denotes expectation when F_{ij} is the underlying distribution function. In the numerical results which follow, the indicated expectations are of course replaced by their Monte Carlo estimates.

The results for sample sizes n=100, 200, 400 and 800 are given in Tables 4, 5, 6 and 7 respectively. In each table p takes on the indicated values 1/50,..., 1/1600. The abbreviations used for the various estimators are: ET (exponential tail), WT(Weibull tail), WML (Weibull maximum likelihood), WP (Weibull percentile), LML (lognormal maximum likelihood), LP (lognormal percentile), and EM (empirical).

Table 4. Summary Statistics for n=100

Measure	Estimator	$\frac{1}{1600}$	$\frac{1}{800}$	$\frac{1}{400}$	$\frac{1}{200}$	$\frac{1}{100}$	$\frac{1}{50}$
$R(\hat{S}(x_p))$	ET 5%	3.24	2.43	1.82	1.36	1.02	0.73
	ET 10%	2.61	2.00	1.57	1.24	0.97	0.72
	ET 25%	2.64	1.85	1.36	1.05	0.85	0.69
	WT 25%	2.38	1.69	1.25	0.96	0.76	0.60
	WT 50%	2.45	1.75	1.28	0.96	0.73	0.55
	WML	0.98	0.90	0.83	0.75	0.67	0.58
	WP 10%	2.49	1.79	1.32	1.02	0.81	0.68
	WP 25%	3.02	2.13	1.54	1.13	0.85	0.64
	LML	22.49	13.06	7.60	4.41	2.52	1.40
	LP 10%	7.24	4.55	2.86	1.80	1.14	0.72
	LP 25%	6.81	4.27	2.69	1.70	1.07	0.68
	EM	4.13	3.02	2.01	1.40	0.99	0.70
$L(\hat{S}(x_p))$	ET 5%	7.71	5.52	3.89	2.66	1.75	1.07
	ET 10%	6.92	4.83	3.30	2.20	1.44	0.91
	ET 25%	9.24	6.34	4.21	2.69	1.64	0.96
	WT 25%	26.08	15.52	9.07	5.14	2.78	1.40
	WT 50%	5.75	4.28	3.11	2.19	1.48	0.95
	WML	4.74	3.73	2.89	2.19	1.61	1.15
	WP 10%	15.81	11.37	7.97	5.41	3.56	2.31
	WP 25%	4.55	3.45	2.57	1.87	1.32	0.91
	LML	2.48	2.10	1.74	1.40	1.07	0.77
	LP 10%	2.62	2.23	1.84	1.48	1.13	0.81
	LP 25%	1.86	1.57	1.30	1.05	0.82	0.63
$R(\hat{x}_p)$	ET 5%	0.84	0.90	0.94	0.84	0.58	0.83
	ET 10%	0.75	0.79	0.80	0.73	0.52	0.74
	ET 25%	0.81	0.82	0.80	0.69	0.48	0.63
	WT 25%	1.59	1.25	1.02	0.78	0.49	0.63
	WT 50%	1.17	1.07	0.95	0.76	0.49	0.63
	WML	0.77	0.81	0.82	0.74	0.53	0.69
	WP 10%	0.94	0.99	1.01	0.93	0.70	1.07
	WP 25%	0.94	0.95	0.94	0.81	0.56	0.77
	LML	23.78	17.20	11.91	7.10	3.45	3.43
	LP 10%	9.29	5.96	3.55	1.99	1.00	1.09
	LP 25%	4.29	3.42	2.58	1.71	0.91	0.99

Table 5. Summary Statistics for n=200

Measure	Estimator	$\frac{1}{1600}$	$\frac{1}{800}$	$\frac{1}{400}$	$\frac{1}{200}$	$\frac{1}{100}$	$\frac{1}{50}$
$R(\hat{S}(x_p))$	ET 5%	1.97	1.54	1.21	0.94	0.71	0.50
	ET 10%	1.67	1.30	1.05	0.85	0.68	0.53
	ET 25%	2.05	1.46	1.08	0.83	0.65	0.51
	WT 25%	1.63	1.21	0.92	0.71	0.55	0.42
	WT 50%	1.64	1.22	0.92	0.70	0.53	0.40
	WML	0.88	0.84	0.79	0.72	0.64	0.54
	WP 10%	1.56	1.17	0.90	0.71	0.57	0.48
	WP 25%	1.74	1.31	1.00	0.78	0.60	0.47
	LML	21.90	12.76	7.44	4.32	2.47	1.36
	LP 10%	4.03	2.62	1.71	1 11	0.72	0.47
	LP 25%	4.54	2.89	1.84	1.17	0.74	0.47
	EM	2.81	1.99	1.37	1.01	0.70	0.48
$L(\hat{S}(x_p))$	ET 5%	4.06	2.80	1.90	1.26	0.83	0.53
	ET 10%	4.59	3.08	2.00	1.27	0.80	0.52
	ET 25%	7.44	5.01	3.23	1.98	1.13	0.62
	WT 25%	4.13	2.96	2.07	1.38	0.88	0.54
	WT 50%	2.34	1.81	1.36	0.99	0.70	0.48
	WML	4.50	3.51	2.69	2.02	1.47	1.03
	WP 10%	3.08	2.40	1.85	1.40	1.07	0.84
	WP 25%	2.59	1.99	1.50	1.10	0.80	0.58
	LML	2.47	2.09	1.73	1.38	1.05	0.75
	LP 10%	1.57	1.32	1.08	0.86	0.65	0.48
	LP 25%	1.46	1.20	0.97	0.75	0.57	0.43
$R(\hat{x}_p)$	ET 5%	0.76	0.79	0.72	0.52	0.73	0.96
	ET 10%	0.73	0.73	0.65	0.46	0.63	0.86
	ET 25%	0.90	0.89	0.78	0.52	0.63	0.72
	WT 25%	1.02	0.91	0.74	0.49	0.61	0.71
	WT 50%	0.85	0.82	0.69	0.47	0.60	0.71
	WML	0.91	0.95	0.87	0.63	0.85	1.02
	WP 10%	0.80	0.82	0.75	0.54	0.78	1.08
	WP 25%	0.79	0.80	0.71	0.51	0.68	0.86
	LML	23.35	17.35	11.13	5.74	5.62	4.77
	LP 10%	2.82	2.17	1.51	0.87	1.00	1.01
	LP 25%	2.85	2.38	1.71	0.97	1.03	1.02

Table 6. Summary Statistics for n=400

Measure	Estimator	$\frac{1}{1600}$	$\frac{1}{800}$	$\frac{1}{400}$	$\frac{1}{200}$	$\frac{1}{100}$	$\frac{1}{50}$
$R(\hat{S}(x_p))$	ET 5%	1.31	1 06	0.87	0.70	0.53	0.37
	ET 10%	1.27	0.98	0.78	0.63	0.51	0.40
	ET 25%	1.80	1.29	0.95	0.72	0.53	0.38
	WT 25%	1.22	0.93	0.72	0.55	0.41	0.29
	WT 50%	1.22	0.93	0.71	0.54	0.40	0.29
	WML	0.86	0.83	0.78	0.71	0.63	0.53
	WP 10%	1.12	0.88	0.70	0.56	0.45	0.38
	WP 25%	1.16	0.90	0.71	0.56	0.44	0.35
	LML	21.50	12.56	7.34	4.26	2.44	1.34
	LP 10%	2.64	1.75	1.16	0.76	0.50	0.33
	LP 25%	3.57	2.25	1.41	0.87	0.53	0.32
	EM	2.07	1.47	1.03	0.73	0.51	0.35
$L(\hat{S}(x_p))$	ET 5%	2.99	1.98	1.29	0.83	0.54	0.35
	ET 10%	3.83	2.49	1.54	0.92	0.55	0.36
	ET 25%	6.85	4.56	2.89	1.71	0.91	0.45
	WT 25%	2.13	1.58	1.13	0.78	0.52	0.33
	WT 50%	1.51	1.16	0.88	0.64	0.45	0.31
	WML	4.48	3.48	2.66	1.98	1.43	0.98
	WP 10%	2.12	1.65	1.26	0.96	0.72	0.56
	WP 25%	2.08	1.57	1.15	0.83	0.59	0.42
	LML	2.46	2.09	1.73	1.38	1.05	0.74
	LP 10%	1.21	1.00	0.81	0.63	0.48	0.35
	LP 25%	1.27	1.02	0.79	0.59	0.42	0.30
$R(\hat{x}_p)$	ET 5%	0.75	0.68	0.48	0.64	0.81	1.09
	ET 10%	0.79	0.70	0.48	0.59	0.70	0.98
	ET 25%	1.08	0.98	0.67	0.77	0.80	0.82
	WT 25%	0.86	0.72	0.48	0.61	0.68	0.78
	WT 50%	0.77	0.69	0.47	0.59	0.67	0.79
	WML	1.13	1.09	0.81	1.03	1.26	1.45
	WP 10%	0.81	0.73	0.51	0.70	0.87	1.22
	WP 25%	0.78	0.69	0.48	0.63	0.76	0.97
	LML	24.27	16.41	8.67	8.75	8.29	6.89
	LP 10%	1.86	1.44	0.86	0.98	1.00	1.05
	LP 25%	2.87	2.24	1.28	1.24	1.16	1.06

Table 7. Summary Statistics for n=800

Measure	Estimator	$\frac{1}{1600}$	$\frac{1}{800}$	$\frac{1}{400}$	$\frac{1}{200}$	$\frac{1}{100}$	$\frac{1}{50}$
$R(\hat{S}(x_p))$	ET 5%	0.92	0.75	0.62	0.51	0.40	0.28
	ET 10%	1.02	0.80	0.62	0.48	0.38	0.32
	ET 25%	1.67	1.20	0.89	0.67	0.47	0.30
	WT 25%	0.90	0.70	0.54	0.40	0.29	0.21
	WT 50%	0.95	0.72	0.54	0.40	0.29	0.21
	WML	0.85	0.82	0.78	0.71	0.62	0.52
	WP 10%	0.82	0.67	0.55	0.44	0.36	0.29
	WP 25%	0.83	0.68	0.55	0.43	0.34	0.27
	LML	21.27	12.45	7.29	4.24	2.43	1.33
	LP 10%	1.71	1.14	0.75	0.50	0.33	0.23
	LP 25%	2.99	1.87	1.15	0.69	0.39	0.23
	EM	1.48	1.04	0.72	0.49	0.35	0.24
$L(\hat{S}(x_p))$	ET 5%	2.19	1.37	0.84	0.52	0.36	0.25
	ET 10%	3.27	2.06	1.22	0.67	0.38	0.27
	ET 25%	6.36	4.19	2.62	1.51	0.76	0.33
	WT 25%	1.10	0.84	0.63	0.45	0.31	0.21
	WT 50%	1.02	0.78	0.58	0.42	0.30	0.21
	WML	4.42	3.43	2.61	1.94	1.39	0.95
	WP 10%	1.48	1.14	0.86	0.65	0.49	0.38
	WP 25%	1.78	1.32	0.95	0.66	0.45	0.32
	LML	2.47	2.09	1.73	1.38	1.05	0.74
	LP 10%	0.89	0.71	0.56	0.43	0.32	0.24
	LP 25%	1.18	0.93	0.70	0.50	0.33	0.22
$R(\hat{x}_p)$	ET 5%	0.61	0.41	0.55	0.68	0.95	1.19
	ET 10%	0.72	0.47	0.60	0.67	0.81	1.05
	ET 25%	1.12	0.77	0.98	1.07	1.09	0.92
	WT 25%	0.61	0.42	0.57	0.67	0.78	0.80
	WT 50%	0.62	0.42	0.56	0.65	0.78	0.82
	WML	1.17	0.86	1.22	1.51	1.84	2.15
	WP 10%	0.66	0.46	0.65	0.80	1.01	1.27
	WP 25%	0.67	0.45	0.61	0.73	0.90	1.09
	LML	22.13	12.00	13.88	13.91	13.24	10.51
	LP 10%	1.25	0.76	0.91	0.95	1.00	0.98
	LP 25%	2.38	1.41	1.59	1.51	1.34	1.03

The results for $R(\hat{x}_p)$ again demonstrate that exponential tail estimators do suprisingly well from an overall viewpoint in estimating the p^{th} quantile x_p for small values of p. Weibull tail estimators do reasonably well when the three criterion $R(\hat{S}(x_p))$, $L(\hat{S}(x_p))$ and $R(\hat{x}_p)$ are considered together. The two-step procedure described in Section 2.3 will hopefully do substantially better.

Recall that $\Gamma(pn)$ is the smallest integer which is greater than or equal to pn, so that $X_{\Gamma(pn)}$ is an empirical estimator of x_p. When F_{ij} is the underlying distribution function, the relative mean square error of this estimator is given by

$$E_{ij}\left(\frac{X_{\Gamma(pn)}-x_p}{x_p}\right)^2 .$$

Monte Carlo estimates of the quantity

$$\left(\frac{1}{20}\sum_{i=1}^{4}\sum_{j=1}^{5} E_{ij}\left(\frac{X_{\Gamma(pn)}-x_p}{x_p}\right)^2\right)^{1/2}$$

as a function of n and p are shown in Table 8. It is interesting to note that for each value of n, this quantity is largest when p=1/n.

Table 8. Summary Statistics for $E\left(\frac{X_{\lceil (pn)}-x_p}{x_p}\right)^2$

n	$\frac{1}{1600}$	$\frac{1}{800}$	$\frac{1}{400}$	$\frac{1}{200}$	$\frac{1}{100}$	$\frac{1}{50}$
100	0.46	0.43	0.44	0.58	0 90	0.48
200	0.40	0.42	0.53	0.81	0.39	0.25
400	0.37	0.45	0.67	0.34	0.24	0.15
800	0.41	0.60	0.32	0.21	0.14	0.11

4. TRANSFORMED EXPONENTIAL TAIL ESTIMATORS

In Section 4.1 transformed exponential estimators will be defined in general and Theorem 4.1 will be obtained, which can be used to verify the existence and uniqueness of such estimators. In Section 4.2 this theorem will be applied to verify the existence and uniqueness of Weibull tail estimators. In Section 4.3 it will be applied to several other specific transformed exponential tail estimators.

4.1 GENERAL FORM

Let x_0 be a positive number and let $T_{x_0}(x)$, $x \geq x_0$, be a continuous strictly increasing function of x such that $T_{x_0}(x_0) = 0$ and $\lim_{x\to\infty} T_{x_0}(x) = \infty$. Let $T_{x_0}^{-1}(y)$, $y \geq 0$, denote the corresponding inverse function. Two examples of such functions T are given by $T_{x_0}(x) = x - x_0$ and $T_{x_0}(x) = \log(x/x_0)$.

In order to motivate the form of the estimators that will be considered, suppose first that the conditional distribution of $T_{x_0}(x)$ given that $X > x_0$ is exponential with mean a. That is, suppose that

$$P(T_{x_0}(X) > y | X > x_0) = e^{-y/a}, \quad y \geq 0.$$

Then

$$P(X>x|X>x_0) = P(T_{x_0}(X) > T_{x_0}(x)|X>x_0) = \exp[-T_{x_0}(x)/a], \quad x \geq x_0,$$

and hence

$$S(x) = S(x_0) \exp[-T_{x_0}(x)/a], \quad x \geq x_0 \quad . \tag{4.1}$$

In general, let $\hat{S}(x_0)$ be an estimate of $S(x_0)$ and let $\hat{a}$ be an estimate of a. Then (4.1) suggests the estimate $\hat{S}(x)$ of $S(x)$ given by

$$\hat{S}(x) = \hat{S}(x_0)\exp[-T_{x_0}(x)/\hat{a}], \qquad x \geq x_0. \tag{4.2}$$

Associated with (4.2) is the corresponding estimate of x_p given by

$$\hat{x}_p = T_{x_0}^{-1}\left(\hat{a}\log\frac{\hat{S}(x_p)}{p}\right), \qquad 0<p\leq\hat{S}(x_0). \tag{4.3}$$

Estimators of the form given by (4.2) and (4.3) are called <u>transformed exponential tail</u> estimators.

Let the values obtained in a random sample of size n from the distribution of X be written in decreasing order as $X_1 > \ldots > X_n$. The number x_0 is allowed to depend on this data; indeed in practice $x_0 = X_{m+1}$ for some integer m such that $2\leq m<n$. The tail probabilities $S(x)$, $x \leq x_0$, can be estimated by

$$\hat{S}(x) = \frac{1}{n}\#\{i:X_i \geq x\}, \qquad 0 \leq x \leq x_0. \tag{4.4}$$

(Here #I denotes the number of points in the set I.) In particular

$$\hat{S}(x_0) = \frac{1}{n}\#\{i:X_i \geq x_0\}. \tag{4.5}$$

Suppose that

$$m = \#\{i:X_i > x_0\} > 2. \tag{4.6}$$

Maximum likelihood considerations suggest estimating a by

$$\hat{a} = \frac{1}{m} \sum_{1}^{m} T_{x_0}(X_i). \tag{4.7}$$

It remains to complete (4.3) by defining $\hat{x}_p$ for $p > \hat{S}(x_0)$. This will be done by setting

$$\hat{x}_p = X_{\lceil(np)}, \qquad \hat{S}(x_0) < p \leq 1 , \tag{4.8}$$

where $\lceil(y)$ denotes the smallest integer which equals or exceeds y. Given x_0 satisfying (4.6) and given T_{x_0}, formulas (4.2)-(4.8) determine completely the estimators of the tail probabilities $S(x)$, $x \geq 0$, and quantiles x_p, $0 < p \leq 1$. It remains to determine x_0 and T_{x_0}.

If X has an exponential distribution with mean a, the conditional distribution of $X - x_0$ given that $X > x_0$ (x_0 here being a fixed positive number) is exponential with mean a. Thus a natural choice of T_{x_0} is given by $T_{x_0}(x) = x - x_0$. For this choice, equations (4.7), (4.2) and (4.3) become respectively

$$\hat{a} = \frac{1}{m} \sum_{1}^{m} (X_i - x_0), \tag{4.9}$$

$$\hat{S}(x) = \hat{S}(x_0)\, e^{-(x-x_0)/\hat{a}}, \qquad x \geq x_0, \tag{4.10}$$

and

$$\hat{x}_p = x_0 + \hat{a} \log \frac{\hat{S}(x_0)}{p}, \qquad 0 < p \leq \hat{S}(x_0) . \tag{4.11}$$

Estimators of the form given by (4.10) and (4.11) are called exponential tail estimators.

If log X has an exponential distribution with mean a (so that X itself has a Pareto distribution), the conditional distribution of $\log(X/x_0)$ given that $X > x_0$ is exponential with mean a. Thus another natural choice of T_{x_0} is given by $T_{x_0}(x) = \log(x/x_0)$. For this choice, equations (4.7), (4.2) and (4.3) become respectively

$$\hat{a} = \frac{1}{m}\sum_{1}^{m} \log \frac{X_i}{x_0}, \tag{4.12}$$

$$\hat{S}(x) = \hat{S}(x_0)\left(\frac{x_0}{x}\right)^{1/\hat{a}}, \qquad x \geq x_0, \tag{4.13}$$

and

$$\hat{x}_p = x_0\left(\frac{\hat{S}(x_0)}{p}\right)^{\hat{a}}, \quad 0 < p < \hat{S}(x_0) \ . \tag{4.14}$$

Estimators of the form given by (4.13) and (4.14) are called Pareto Tail estimators. Such estimators were considered by Hill [7].

If X^b has an exponential distribution with mean a for some $b > 0$ (so that X itself has a Weibull distribution), the conditional distribution of $X^b - x_0^b$ given that $X > x_0$ is exponential with mean a. Thus a third natural choice of T_{x_0} is given by $T_{x_0}(x) = x^b - x_0^b$. This choice will be treated in Section 2.2.

The last example suggests choosing a family of transformations $T_{x_0}(x,b)$, $x \geq x_0$, where b ranges over some interval and is chosen adaptively from the sample data. This approach, in comparison to that of choosing a fixed transformation, is applicable to a wider range of the unknown underlying distribution, but it requires a larger sample size to be reliable.

In implementing this approach, there is a variety of reasonable ways to choose a value $\hat{b}$ for b based on the data X_i, $1 \leq i \leq m$, where m is given as before by (4.6). Given $\hat{b}$, it is natural to choose $\hat{a}$ as in (4.7), so that

$$\hat{a} = \frac{1}{m} \sum_{1}^{m} T_{x_0}(X_i, \hat{b}). \tag{4.15}$$

Hill [4] suggested choosing $\hat{b}$ by the maximum likelihood method. According to this method $\hat{b}$ should be chosen to maximize

$$\prod_{1}^{n} \frac{t_{x_0}(X_i, \hat{b})}{\hat{a}} \exp[-T_{x_0}(X_i, \hat{b})/\hat{a}],$$

where $\hat{a}$ is defined in terms of $\hat{b}$ according to (4.15) and $t_{x_0}(x,b) = \partial T_{x_0}(x,b)/\partial x$. In a different context, Box and Cox [2] had earlier suggested choosing the parameters of a transformation by the maximum likelihood method. But there is no compelling reason for this choice unless the corresponding theoretical model is really believed to be valid.

Another approach to choosing $\hat{b}$ is to make some feature of the empirical distribution of $T_{x_0}(X_i, \hat{b})/\hat{a}$, $1 \leq i \leq m$, match up with the corresponding feature of the standard exponential distribution. The work of Pickands [12] suggests matching up some fixed quantile. The test for exponentiality due to Shapiro and Wilk [14] (see also Stephens [15]) suggests making the empirical second moment of $T_{x_0}(X_i, \hat{b})/\hat{a}$, $1 \leq i \leq m$, equal to the second moment of the standard exponential distribution, namely 2. According to this criterion, $\hat{b}$ should be chosen to satisfy the equation $U(\hat{b}) = 2$, where

$$U(b) = \frac{\frac{1}{m}\sum_{1}^{m} T_{x_0}^2(X_i, b)}{\left(\frac{1}{m}\sum_{1}^{m} T_{x_0}(X_i, b)\right)^2} .$$

This is the criterion which will be studied in the present report.

In using this criterion it is important to determine, for a given family of transformations, when a solution to $U(\hat{b}) = 2$ exists and when it is unique. That this determination can be made for a variety of natural families of transformations is one of the main justifications for the criterion under consideration.

In analyzing the equation $U(\hat{b}) = 2$ it is convenient to define U in a more general manner. To this end, think of X_i, $1 \le i \le m$, as fixed numbers and let $\tilde{X}$ denote a random variable whose distribution is the same as the empirical distribution of X_i, $1 \le i \le m$. Also set $T(x, b) = T_{\tilde{x}_0}(x, b)$, temporarily ignoring the dependence on x_0, and set $T = T(\tilde{X}, b)$. Then $U(b) = ET^2/(ET)^2$.

More generally, let X denote an arbitrary (i.e., not necessarily having distribution function F) nonconstant random variable and let I be an interval such that $P(X \in I) = 1$. Let J be an open interval and let $T(x,b)$ be a strictly increasing, strictly positive continuous function of x on $I \times J$ such that $\partial T(x,b)/\partial b$ exists and is continuous in x on $I \times J$. Set $T = T(X,b)$ and $T' = \partial T(X,b)/\partial b$. Suppose that T^2, T' and TT' have finite expectations and that $dET/db = ET'$ and $dET^2/db = ETT'$ (i.e., that differentiation and expectation can be interchanged here). Define U as before by $U(b) = ET^2/(ET)^2$, $b \in J$. If the interval J is not compact, but U is monotonic

on this interval, let it be defined by continuity as a finite or an infinite number at the missing finite for infinite end points of J.

The following theorem provides a sufficient condition for there to be at most one point $\hat{b} \in J$ such that $U(\hat{b}) = 2$.

Theorem 4.1. Suppose the above conditions hold and that $\partial \log T(x,b)/\partial b$ is strictly increasing (decreasing) in x on I x J. Then U is continuous and strictly increasing (decreasing) and has strictly positive (negative) derivative on J.

Proof. Observe that

$$U' = \frac{2}{(ET)^2} E(T(T' - T \frac{ET'}{ET})).$$

Suppose, say, that

$$\frac{\partial}{\partial b} \log T(x,b) = \frac{\frac{\partial}{\partial b} T(x,b)}{T(x,b)}$$

is strictly increasing in x on I x J. (A similar argument works when this function is strictly decreasing.) It suffices to prove that

$$ET(T' - T \frac{ET'}{ET})) > 0 \quad \text{on} \quad J.$$

Choose $b \in J$. Since X is a nonconstant I-valued random variable, there is an $x_0 \in J$ such that

$$\frac{\partial}{\partial b} T(x,b) > \frac{ET'}{ET} T(x,b), \qquad x \in I \cap (x_0, \infty),$$

and

$$\frac{\partial}{\partial b} T(x,b) < \frac{ET'}{ET} T(x,b), \qquad x \in I \cap (-\infty, x_0).$$

Now

$$E(T(T' - T\frac{ET'}{ET}); X > x_0) \geq T(x_0,b)E(T' - T\frac{ET'}{ET}; X > x_0)$$

and

$$E(T(T' - T\frac{ET'}{ET}); X < x_0) \geq T(x_0,b)E(T' - T\frac{ET'}{ET}; X < x_0).$$

(Here $E(Z;\ A) = E(Z\ I_A)$, where I_A equals 1 or 0 according as the event A does or does not occur.) Note that strict inequality holds in the first (or second) of these two inequalities unless $P(X > x_0) = 0$ (or $P(X < x_0) = 0$). Since X is a nonconstant random variable

$$E(T(T - T'\frac{ET'}{ET})) > 0$$

as desired.

4.2 WEIBULL TAIL ESTIMATORS

Theorem 4.1 will now be applied to the family of transformations given by $T_{x_0}(x,b) = x^b - x_0^b$. Here X is an arbitrary random variable and U is defined in terms of T as before.

<u>Theorem 4.2</u>. Let $0 \leq x_0 < \infty$ and let X be a nonconstant random variable such that $P(X > x_0) = 1$ and $b_0 = \sup[b: EX^{2b} < \infty] > 0$. Set $T(x,b) = x^b - x_0^b$ for $x > x_0$ and $0 < b < b_0$. Then U is continuous and strictly increasing on $(0,\infty)$ and

$$U(0) = \frac{E \log^2 \frac{X}{x_0}}{\left(E \log \frac{X}{x_0}\right)^2} \qquad \text{if } x_0>0, \tag{4.16}$$

$$= 1 \qquad \text{if } x_0 = 0.$$

Suppose $b_0 = \infty$. If there is an $M>0$ such that $P(X<M)<P(X\leq M) = 1$, then $U(\infty) = 1/P(X=M)$. Otherwise $U(\infty) = \infty$.

Theorem 4.2 will be proven at the end of this subsection. For $T_{x_0}(x,b) = x^b-x_0^b$, $x>x_0>0$, U is given by

$$U(b) = \frac{\frac{1}{m}\sum_1^m (X_i^b-x_0^b)^2}{\left(\frac{1}{m}\sum_1^m (X_i^b-x_0^b)\right)^2} .$$

By Theorem 4.2 (with the distribution of X being the empirical distribution of $X_1,\ldots, X_m$), U is continuous and strictly increasing on $(0,\infty)$, $U(\infty)=m>2$ and

$$U(0) = \frac{\frac{1}{m}\sum_1^m \log^2 \frac{X_i}{x_0}}{\left(\frac{1}{m}\sum_1^m \log \frac{X_i}{x_0}\right)^2} .$$

In particular there is at most one positive solution $\hat{b}$ to the equation $U(\hat{b})=2$. There is such a solution if and only if

$$\frac{\frac{1}{m}\sum_{1}^{m}\log^2\frac{X_i}{X_0}}{\left(\frac{1}{m}\sum_{1}^{m}\log\frac{X_i}{X_0}\right)^2} < 2. \tag{4.17}$$

Suppose (4.17) holds. Then (4.15), (4.2) and (4.3) become respectively

$$\hat{a} = \frac{1}{m}\sum_{1}^{m}\left(X_i^{\hat{b}} - x_0^{\hat{b}}\right), \tag{4.18}$$

$$\hat{S}(x) = \hat{S}(x_0)\, e^{-(x^{\hat{b}}-x_0^{\hat{b}})/\hat{a}}, \qquad x \geq x_0, \tag{4.19}$$

and

$$\hat{x}_p = \left(x_0^{\hat{b}} + \hat{a}\log\frac{\hat{S}(x_0)}{p}\right)^{1/\hat{b}}, \qquad 0 < p \leq \hat{S}(x_0) . \tag{4.20}$$

Estimators of the form given by (4.19) and (4.20) are called _Weibull tail_ estimators. Some of their asymptotic properties will be obtained in Section 5.

Proof of Theorem 4.2. It will first be shown that $\partial \log T(x,b)/\partial b$ is strictly increasing in x for $x > x_0$ and $b > 0$. If $x_0 = 0$, then $\partial \log T(x,b)/\partial b = \log x$ and hence the desired result is valid.

Suppose, next that $x_0 > 0$. Then

$$\frac{\partial}{\partial b}\log T(x,b) = \frac{x^b \log x - x_0^b \log x_0}{x^b - x_0^b}$$

$$= \frac{1}{b}\,\frac{z\log z - z_0 \log z_0}{z - z_0}$$

$$= \frac{1}{b}\left(\frac{\frac{z}{z_0}\log\frac{z}{z_0}}{\frac{z}{z_0}-1} + \log z.\right),$$

where $z = x^b > x_0^b = z_0$ for $x > x_0$. Thus is suffices to show that $z \log z/(z-1)$ is strictly increasing in z for $z>1$. Now

$$\frac{d}{dz}\left(\frac{z \log z}{z-1}\right) = \frac{\log z + 1}{z-1} - \frac{z \log z}{(z-1)^2},$$

so it suffices to show that

$$(z-1)(\log z + 1) - z \log z > 0 \qquad \text{for } z>1$$

or, equivalently, that $\log z < z-1$ for $z>1$. The last inequality is obviously true since $d\log z/dz = z^{-1} < 1$ for $z>1$.

Therefore, in general, $\partial \log T(x,b)/\partial b$ is strictly increasing in x for $x>0$ and $b>0$. Thus by Theorem 4.1, U is continuous and strictly increasing on $(0,b_0)$.

If $x_0=0$, then $\lim_{b\to 0} EX^b = \lim_{b\to 0} EX^{2b} = 1$ and hence

$$U(0) = \lim_{b\to 0} \frac{EX^{2b}}{(EX^b)^2} = 1.$$

Consequently (4.16) holds when $x_0=0$.

Suppose instead that $x_0>0$. Since

$$U(b) = \frac{E\left(\left(\frac{X}{x_0}\right)^b - 1\right)^2}{\left[E\left(\left(\frac{X}{x_0}\right)^b - 1\right)\right]^2},$$

in verifying (4.16) it can be assumed that $x_0=1$ and hence that $P(X>1)=1$. Observe that if $x>1$, then

$$\lim_{b\to 0} \frac{x^b-1}{b} = \log x$$

and

$$\frac{x^b-1}{b} = \int_1^x z^{b-1}dz \leq \int_1^x z^{\varepsilon-1}\, dz = \frac{x^{\varepsilon}-1}{\varepsilon} \qquad \text{for } 0<b\leq\varepsilon.$$

Thus by the dominated convergence theorem

$$\lim_{b\to 0} E\left(\frac{X^b-1}{b}\right) = E\log X$$

and

$$\lim_{b\to 0} E\left[\left(\frac{X^b-1}{b}\right)^2\right] = E\log^2 X.$$

Consequently (4.16) holds when $x_0 =1$.

Suppose $b_0=\infty$. Suppose also that (i) there is an $M>x_0$ such that $P(X<M) < P(X\leq M)=1$. Then

$$\lim_{b\to\infty} E\left(\left(\frac{X}{M}\right)^b - \left(\frac{x_0}{M}\right)^b\right) = \lim_{b\to\infty} E\left(\left(\frac{X}{M}\right)^b ;\, X<M\right) + P(X=M)$$

$$= P(X=M)$$

and similarly

$$\lim_{b\to\infty} E\left[\left(\left(\frac{X}{M}\right)^b - \left(\frac{x_0}{M}\right)^b\right)^2\right] = P(X=M) .$$

Therefore

$$U(\infty) = \lim_{b\to\infty} \frac{E\left[\left(\left(\frac{X}{M}\right)^b - \left(\frac{x_0}{M}\right)^b\right)^2\right]}{\left[E\left(\left(\frac{X}{M}\right)^b - \left(\frac{x_0}{M}\right)^b\right)\right]^2} = \frac{1}{P(X=M)} .$$

Suppose instead that (i) fails to hold. Choose $\varepsilon>0$ and let $M>x_0$ be such that $0<P(X\geq M) \leq \varepsilon$. Then

$$\frac{E(X^b - x_0^b;\ X\geq M)}{E(X^b - x_0^b)} = 1$$

and

$$\lim_{b\to\infty} \frac{E((X^b - x_0^b)^2;\ X\geq M)}{E(X^b - x_0^b)^2)} = 1.$$

By Schwarz's inequality

$$E((X^b - x_0^b)^2;\ X\geq M) \geq \frac{[E((X^b-x_0^b);\ X\geq M)]^2}{P(X\geq M)}$$

$$\geq \frac{[E((X^b-x_0^b);\ X\geq M)]^2}{\varepsilon} .$$

Consequently

$$U(\infty) = \lim_{b\to\infty} \frac{E((X^b-x_0^b)^2;\ X\geq M)}{[E(X^b-x_0^b;\ X\geq M]^2} \geq \frac{1}{\varepsilon} .$$

Therefore $U(\infty) = \infty$, which completes the proof of the theorem.

4.3 OTHER SPECIAL CASES

Theorem 4.1 will now be applied to several other elementary families of transformations. The first such family covers situations in which (4.17) fails to hold.

Theorem 4.3 Let $x_0>0$ and let X be a nonconstant random variable such that $P(X>x_0)=1$ and $E\log^2 X<\infty$. Set $T(x,b)=\log(1+b\frac{x-x_0}{x_0})$ for $x>x_0$ and $b>0$. Then U is continuous and strictly decreasing on $(0,\infty)$. If $EX^2<\infty$, then $U(0)=E(X-x_0)^2/(E(X-x_0))^2$.

Proof Observe that

$$\frac{\partial}{\partial b}\log T(x,b) = \frac{\frac{x-x_0}{x_0}}{(1+b\frac{x-x_0}{x_0})\log(1+b\frac{x-x_0}{x_0})}\ .$$

Thus to prove that $\partial\log T(x,b)/\partial b$ is strictly decreasing in x for $x>x_0$, it suffices to show that

$$\frac{(1+x)\log(1+x)}{x}$$

is strictly increasing in x for $x>0$. But this result was verified in the proof of Theorem 4.2. The conclusions of the present theorem now follow easily from Theorem 4.1 and the dominated convergence theorem.

Consider the family of transformations given by

$$T_{x_0}(x,b) = \log(1+b\,\frac{x-x_0}{x_0}) , \quad x>x_0 ,$$

where $b>0$. Then

$$U(b) = \frac{\frac{1}{m}\sum_1^m \log^2 (1+b\,\frac{X_i-x_0}{x_0})}{(\frac{1}{m}\sum_1^m \log (1+b\,\frac{X_i-x_0}{x_0}))^2} .$$

By Theorem 4.3, U is continuous and strictly decreasing on $(0,\infty)$, $U(\infty)=1$ and

$$U(0) = \frac{\frac{1}{m}\sum_1^m (X_i-x_0)^2}{(\frac{1}{m}\sum_1^m (X_i-x_0))^2} .$$

Thus there is at most one positive solution to the equation $U(\hat{b})=2$, and there is such a solution if and only if

$$\frac{\frac{1}{m}\sum_1^m (X_i-x_0)^2}{(\frac{1}{m}\sum_1^m (X_i-x_0))^2} > 2 . \tag{4.21}$$

Observe that the left side of (4.21) is larger than the left side of (4.17), so that (4.17) and (4.21) cannot simultaneously fail. Suppose (4.21) holds. Then (4.15), (4.2) and (4.3) become respectively

$$\hat{a} = \frac{1}{m}\sum_{1}^{m} \log\left(1+\hat{b}\,\frac{X_i - x_0}{x_0}\right), \tag{4.22}$$

$$\hat{S}(x) = \frac{\hat{S}(x_0)}{\left(1+\hat{b}\,\frac{x-x_0}{x_0}\right)^{1/\hat{a}}}\,, \qquad x \geq x_0, \tag{4.23}$$

and

$$\hat{x}_p = x_0\left(\frac{1}{\hat{b}}\left(\frac{\hat{S}(x_0)}{p}\right)^{\hat{a}} + 1 - \frac{1}{\hat{b}}\right), \qquad 0<p \leq \hat{S}(x_0)\ . \tag{4.24}$$

Estimators of the form given by (4.23) and (4.24) are called Pareto tail estimators. When $\hat{b}=1$ they reduce to the forms given by (4.13) and (4.14).

Suppose (4.17) and (4.21) both hold; i.e., that

$$\frac{\frac{1}{m}\sum_{1}^{m} \log^2 \frac{X_i}{x_0}}{\left(\frac{1}{m}\sum_{1}^{m} \log \frac{X_i}{x_0}\right)^2} < 2 < \frac{\frac{1}{m}\sum_{1}^{m} (X_i - x_0)^2}{\left(\frac{1}{m}\sum_{1}^{m} (X_i - x_0)\right)^2}\ . \tag{4.25}$$

Then there is a Weibull tail estimator for a unique choice of $\hat{b}$, with $0<\hat{b}<1$; and there is a Pareto tail estimator for a unique choice of $\hat{b}$, again with $0<\hat{b}<1$. When (4.25) holds it seems reasonable to choose the method which corresponds to the larger value of $\hat{b}$.

There are some other elementary families of transformations for which Theorem 4.1 is applicable.

Theorem 4.4. Let $c \neq 1$ be a positive constant. Let $x_0 \geq 0$ and let X be a nonconstant random variable such that $P(X > x_0) = 1$, $EX^{2c} < \infty$ and $EX^2 < \infty$. Set $T(x,b) = (x+b)^c - (x_0+b)^c$ for $x > x_0$ and $b \geq x_0$. Then U is continuous on $[-x_0, \infty)$,

$$U(-x_0) = \frac{E(X-x_0)^{2c}}{(E(X-x_0)^c)^2} ,$$

and

$$U(\infty) = \frac{E(X-x_0)^2}{(E(X-x_0))^2} .$$

Also U is strictly increasing or strictly decreasing on $[-x_0, \infty)$ according as $0<c<1$ or $c>1$.

Proof. It is clear that U is continuous on $[-x_0, \infty)$ and that $U(-x_0)$ has the indicated value. The value for $U(\infty)$ is easily verified by using the dominated convergence theorem and the formula

$$b((1+ \frac{x}{b})^c - 1) = \int_0^x c(1+ \frac{y}{b})^{c-1} dy.$$

Observe that

$$\frac{\partial}{\partial b} \log T(x,b) = \frac{c}{b} \frac{(\frac{x}{b}+1)^{c-1} - 1}{(\frac{x}{b}+1)^c - 1} .$$

A simple computation yields that

$$\frac{d}{dx}\left(\frac{x^{c-1}-1}{x^{c}-1}\right) = \frac{x^{c-2}}{(x^{c}-1)^{2}}\,(-x^{c} + cx - (c-1)) \ .$$

It follows easily from this the formula

$$\frac{d}{dx}\,(-x^{c} + cx - (c-1)) = c(1- x^{c-1})$$

that $(x^{c-1}-1)/(x^{c}-1)$ is strictly increasing or strictly decreasing on $[1,\infty)$ according as $0<c<1$ or $c>1$. The last conclustion of the Theorem now follows from Theorem 2.1.

<u>Theorem 4.5</u>. Let $x_0 \geq 0$ and let X be a nonconstant random variable such that $P(X>x_0)=1$ and $EX^2<\infty$. Set $T(x,b) = \log(x+b)-\log(x_0+b)$ for $x>x_0$ and $b>-x_0$. Then U is continuous and strictly increasing on $(-x_0,\infty)$, $U(-x_0) = 1$ and $U(\infty) = E(X-x_0)^2/E(X-x_0))^2$.

The proof of this theorem is similar to that of Theorem 4.3 and left to the reader. It is also straightforward and left to the reader to analyze the transformed exponential tail estimators corresponding to the transformations defined in the statements of Theorems 4.4 and 4.5.

5. ASYMPTOTIC BEHAVIOR OF WEIBULL TAIL ESTIMATORS

In this section the asymptotic behavior of Weibull tail estimators will be studied in a modified setting where the analysis is less complicated. The results will be described in Section 5.1 and proven in Section 5.2.

5.1 DISCUSSION OF RESULTS

Let X have distribution F as before. Given $x_0>0$, let Y be distributed as the conditional distribution of $X-x_0$ given that $X>x_0$. Then

$$P(Y>y) = P(X-x_0>y \mid X>x_0) = \frac{S(x_0+y)}{S(x_0)} , \quad y \geq 0 \ .$$

Given a positive integer $m>2$, let $Y_1,\ldots,Y_m$ be independent variables each having the same distribution as Y. Since F is a continuous distribution function, $Y_1,\ldots,Y_m$ are distinct (with probability one).

Set

$$\hat{U}(b) = \frac{\frac{1}{m}\sum_1^m ((x_0+Y_i)^b - x_0^b)^2}{\left[\frac{1}{m}\sum_1^m ((x_0+x_i)^b - x_0^b\right]^2} , \quad b>0 \ .$$

By Theorem 4.2, $\hat{U}$ is continuous and strictly increasing on $(0,\infty)$, $U(\infty) = m>2$, and

$$\hat{U}(0) = \frac{\frac{1}{m}\sum_1^m \log^2\left(\frac{x_0+y_i}{x_0}\right)}{\left(\frac{1}{m}\sum_1^m \log \frac{x_0+y_i}{x_0}\right)^2} \ .$$

Thus the equation $\hat{U}(\hat{b}) = 2$ has at most one positive solution $\hat{b}$. It does have such a solution if and only if

$$\frac{\frac{1}{m}\sum_1^m \log^2\left(\frac{x_0+Y_i}{x_0}\right)}{\left(\frac{1}{m}\sum_1^m \log\left(\frac{x_0+Y_i}{x_0}\right)\right)^2} < 2 \quad . \tag{5.1}$$

When (5.1) holds set

$$\hat{a} = \frac{1}{m}\sum_1^m ((x_0+Y_i)^{\hat{b}} - x_0^{\hat{b}}) \quad .$$

According to (4.19)

$$\hat{S}(x_0+y) = \hat{S}(x_0)\, e^{-((x_0+Y)^{\hat{b}} - x_0^{\hat{b}})/\hat{a}} \quad , \quad y \geq 0 \quad .$$

It is convenient to set

$$\hat{s}(y) = \log \frac{\hat{S}(x_0)}{\hat{S}(x_0+y)} \quad , \quad y \geq 0 \quad ,$$

so that

$$\hat{s}(y) = \frac{(x_0+y)^{\hat{b}} - x_0^{\hat{b}}}{\hat{a}} \quad , \quad y \geq 0 \quad .$$

A problem of obvious interest is to determine the asymptotic distribution of $\hat{s}(y)$ as an estimator of $\log(S(x_0)/S(x_0+y))$ as $m \to \infty$ and $x_0 \to \infty$ in some manner. To do so, it is necessary to determine both the asymptotic variance and asymptotic bias of $\hat{s}(y)$. The asymptotic variance of $\hat{s}(y)$ will

be studied here as $m \to \infty$ for fixed x_0 and then $x_0 \to \infty$. (The asymptotic bias appears to be amenable only to numerical analysis in specific cases.)

As $m \to \infty$ for fixed x_0, the random quantities $\hat{U}$, $\hat{b}$, $\hat{a}$ and $\hat{s}$ converge in probability to certain quantities U, b, a and s, which will now be defined. Set

$$U(b) = \frac{E((x_0+Y)^b - x_0^b)^2}{[E((x_0+Y)^b - x_0^b)]^2}, \quad b>0 \quad .$$

Assume that all moments of Y are finite. Then by Theorem 4.2, U is continuous and strictly increasing on $(0,\infty)$, $U(\infty)=\infty$ and

$$U(0) = \frac{E \log^2 \frac{x_0+Y}{x_0}}{E \log\left(\frac{x_0+Y}{x_0}\right)^2} \quad .$$

Thus the equation $U(b)=2$ has at most one positive solution b. It does have such a solution if and only if

$$\frac{E \log^2 \frac{x_0+Y}{x_0}}{\left(E \log \frac{x_0+Y}{x_0}\right)^2} < 2 \quad . \tag{5.2}$$

When (5.2) holds, set

$$a = E((x_0+Y)^b - x_0^b)$$

and

$$s(y) = \frac{(x_0+y)^b - x_0^b}{a} \qquad y \geq 0 \quad .$$

Assume that (5.2) holds for all sufficiently large values of x_0 and also for the specific value of x_0 under consideration. Then by the weak law of large numbers, (5.1) holds with probability approaching one as $m \to \infty$ for fixed x_0.

The random variables $\hat{b}$, $\hat{a}$ and $s(y)$ converge in probability to b, a and $s(y)$ respectively as $m \to \infty$ for fixed x_0. Moreover there are positive numbers $\sigma_{\hat{b}}$, $\sigma_{\hat{a}}$ and $\sigma_{\hat{s}(y)}$ (depending on x_0) such that as $m \to \infty$ for fixed x_0

$$\mathcal{L}(\sqrt{m}\,(\hat{b}-b)) \to N(0,\sigma^2_{\hat{b}}) \quad ,$$

$$\mathcal{L}(\sqrt{m}\,(a-a)) \to N(0,\sigma^2_{\hat{a}})$$

and

$$\mathcal{L}(\sqrt{m}(\hat{s}(y) - s(y))) \to N(0,\sigma^2_{\hat{s}(y)}) \quad .$$

Given $z>0$, let y be defined in terms of z (and x_0) by

$$s(y) = \frac{(x_0+y)^b - x_0^b}{a} = z \quad .$$

Also set

$$Z = \frac{(x_0+Y)^b - x_0^b}{a}$$

and note that $EZ=1$.

The asymptotic behavior of $\sigma_{\hat{b}}$, $\sigma_{\hat{a}}$ and $\sigma_{\hat{s}(y(z))}$ as $x_0 \to \infty$ will now be determined.

Theorem 5.1. Suppose that

$$\frac{a}{x_0^b} = E(1 + Y/x_0)^b - 1 \to 0 \quad \text{as } x_0 \to \infty .$$

Suppose also that as $x_0 \to \infty$ the distribution of Z converges to the exponential distribution with mean one and that EZ^4 converges to the fourth moment of the limiting distribution (which equals 24). Then as $x_0 \to \infty$

$$\frac{a}{b\, x_0^b}\, \sigma_{\hat{b}} \to 1 \quad , \tag{5.3}$$

$$\frac{1}{(a^2 + (b\, x_0^b\, (\log x_0 + 1/b)\,)^2)^{1/2}}\, \sigma_{\hat{a}} \to 1 \quad , \tag{5.4}$$

and

$$\sigma_{\hat{s}(y(z))} \to z^2((z/2 - 1)^2 + 1) \quad . \tag{5.5}$$

By itself, Theorem 5.1 doesn't provide much justification for using Weibull tail estimators. In order to obtain some such justification, suppose that the Weibull tail model is exactly valid for the value of x_0 under consideration (and hence for all larger values of x_0); that is, suppose that

$$P(Y \geq y) = \exp\left[-\left(\left(\frac{x_0 + y}{c}\right)^b - \left(\frac{x_0}{c}\right)^b\right)\right], \quad y \geq 0 \quad ,$$

where $c = a^{1/b}$. Then above assumptions, including those of Theorem 5.1, are clearly satisfied. Also $\hat{s}(y)$ is a consistent estimator of

$$s(y) = \log \frac{S(x_0)}{S(x_0+y)}$$

and as $m\to\infty$ for fixed x_0

$$\mathcal{L}\left(\sqrt{m}\left(\hat{s}(y) - \log \frac{S(x_0)}{S(x_0+y)}\right)\right) \to N\left(0,\sigma^2_{\hat{s}(Y)}\right) .$$

Let $\bar{b}$, $\bar{c}$ and $\bar{s}(y)$ be consistent maximum likelihood estimators of b, c and s(y) based on the random sample $Y_1,\ldots,Y_m$. Then there are positive numbers $\sigma_{\bar{b}}$, $\sigma_{\bar{c}}$ and $\sigma_{\bar{s}(y)}$ (depending on x_0) such that as $m\to\infty$ for fixed x_0

$$\mathcal{L}(\sqrt{m}\ (\bar{b}-b)) \to N\left(0,\sigma^2_{\bar{b}}\right) ,$$

$$\mathcal{L}(\sqrt{m}\ (\bar{c}-c)) \to N\left(0,\sigma^2_{\bar{c}}\right)$$

and

$$\mathcal{L}(\sqrt{m}\ \bar{s}(y)-s(y))) \to N\left(0,\sigma^2_{\bar{s}(y)}\right)$$

Theorem 5.2. Suppose the Weibull tail model is exactly valid for some x_0. Then as $x_0\to\infty$

$$\frac{a}{b\ x_0^b}\,\sigma_{\bar{b}} \to 1 \tag{5.6}$$

and

$$\sigma^2_{\bar{s}(y(z))} \to z^2((z/2 - 1)^2 + 1) \quad . \tag{5.7}$$

Thus $\hat{b}$ and $\hat{s}(y(z))$ are asymptotically efficient estimators of b and $s(y(z))$ respectively as first $m\to\infty$ and then $x_0\to\infty$.

5.2 PROOFS

Let $0_p(1)(o_p(1))$ denote a random variable which is bounded (converges to zero in probability) as $m\to\infty$. Given a sequence $\{c_m\}$ of positive constants, let $0_p(c_m)(o_p(c_m))$ denote a random variable which is of the form $c_m 0_p(1)(c_m o_p(1))$.

Define $Z_1, Z_2 \ldots$ by

$$Z_i = \frac{(x_0+Y_i)^b - x_0^b}{a} \quad .$$

Then the Z_i's are independent random variables each having the same distribution as

$$Z = \frac{(x_0+Y)^b - x_0^b}{a} \quad .$$

Let μ_k denote the k^{th} moment of Z (which depends on x_0). Then $\mu_1=1$.

Observe that

$$\hat{U}(b) = \frac{\frac{1}{m}\sum_1^m Z_i^2}{\left(\frac{1}{m}\sum_1^m Z_i\right)^2} = \frac{\mu_2 + \frac{1}{m}\sum_1^m (Z_i^2 - \mu_2)}{(1 + \frac{1}{m}\sum_1^m (Z_i - 1))^2} \quad .$$

By the central limit theorem

$$\frac{1}{m}\sum_{1}^{m}(Z_i-1) = O_p(1/m^{1/2})$$

and

$$\frac{1}{m}\sum_{1}^{m}(Z_i^2 - \mu_2) = O_p(1/m^{1/2}) \quad .$$

Thus

$$\hat{U}(b) = \frac{\mu_2 + \frac{1}{m}\sum_{1}^{m}(Z_i^2 - \mu_2)}{1 + \frac{2}{m}\sum_{1}^{m}(Z_i-1) + O_p\left(\frac{1}{m}\right)}$$

$$= \left(\mu_2 + \frac{1}{m}\sum_{1}^{m}(Z_i^2 - \mu_2)\right)\left(1 - \frac{2}{m}\sum_{1}^{m}(Z_i-1)\right) + O_p\left(\frac{1}{m}\right)$$

and consequently

$$\hat{U}(b) - U(b) = \frac{1}{m}\sum_{1}^{m}(Z_i^2 - 2\mu_2 Z_i + \mu_2) + O_p\left(\frac{1}{m}\right) .$$

Therefore by the central limit theorem

$$\mathcal{L}\left(\sqrt{m}(\hat{U}(b) - U(b))\right) \to N(0, \mu_4 - 4\mu_2\mu_3 + 4\mu_2^3 - \mu_2^2) \quad .$$

Now $\hat{b}$ is a consistent estimator of b. To see this, let $0<b_1<b<b_2<\infty$. Then $U(b_1)<U(b) = 2<U(b_2)$ according to Theorem 4.2. By the weak law of large numbers, $\hat{U}(b_1) \to U(b_1) < 2$ in probability and $\hat{U}(b_2) \to U(b_2) > 2$ as $m\to\infty$. Since $\hat{U}$ is monotonic, the probability that the inequality $b_1<\hat{b}<b_2$

holds converges to one as $m\to\infty$. Thus $\hat{b}$ converges in probability to b as $m\to\infty$, as desired.

By the mean value theorem of calculus

$$\hat{U}(\hat{b}) = \hat{U}(b) + \hat{U}'(b_0)(\hat{b}-b) = 2,$$

where b_0 is between b and $\hat{b}$. Thus

$$\hat{b} = b - \frac{\hat{U}(b) - 2}{\hat{U}'(b_0)} \quad .$$

Since $\hat{b}$ converges in probability to b as $m\to\infty$, so does b_0. It follows by a direct computation that $\hat{U}'(b)$ converges in probability to $U'(b)$ as $m\to\infty$ and that

$$\max_{b_1\le b_3\le b_2} |U''(b_3)| = O_p(1) \qquad \text{for } 0<b_1<b<b_2<\infty \; .$$

Consequently $\hat{U}'(b_0)$ converges in probability to $U'(b)$ as $m\to\infty$. Therefore

$$\hat{b}-b = -\frac{\hat{U}(b)-2}{U'(b_0)} + o_p\left(\frac{1}{m^{1/2}}\right)$$

and

$$\mathcal{L}(\sqrt{m}\;(\hat{b}-b)) \to N\left(0,\sigma^2_{\hat{b}}\right) \quad ,$$

where

$$\sigma^2_{\hat{b}} = \frac{1}{(U'(b))^2} \; (\mu_4 - 4\mu_2\mu_3 + 4\mu_2^3 - \mu_2^2) \; .$$

Define functions $a(\cdot)$ and $\hat{a}(\cdot)$ by

$$a(b) = E((x_0+Y)^b - x_0^b), \qquad b>0,$$

and

$$\hat{a}(b) = \frac{1}{m}\sum_1^m ((x_0+Y_i)^b - x_0^b), \qquad b>0 \quad .$$

Then

$$\hat{a}(b) - a = \frac{a}{m}\sum_1^m (Z_i-1) \ .$$

By the mean value theorem

$$\hat{a} = \hat{a}(\hat{b}) = \hat{a}(b) + \hat{a}'(b_0)(\hat{b}-b),$$

where b_0 is again between b and $\hat{b}$. It follows as in the previous paragraph that $\hat{a}'(b_0)$ converges in probability to $a'(b)$ as $m\to\infty$. Thus

$$\begin{aligned}\hat{a}-a &= \hat{a}(b) - a + a'(b)(\hat{b}-b) + o_p\left(\frac{1}{m^{1/2}}\right) \\ &= \frac{1}{m}\sum_1^m [a(Z_i-1) + \frac{a'(b)}{U'(b)}(Z_i^2 - 2\mu_2 Z_i + 1)] + o_p\left(\frac{1}{m^{1/2}}\right) \quad .\end{aligned}$$

Therefore by the central limit theorem

$$\mathcal{L}(\sqrt{m}\,(\hat{a}-a)) \to N\left(0,\sigma_{\hat{a}}^2\right) ,$$

where

$$\sigma_{\hat{a}}^2 = a^2(\mu_2-1) + \left(\frac{a'(b)}{U'(b)}\right)^2 (\mu_4 - 4\mu_2\mu_3 + 4\mu_2^3 - \mu_2^2)$$

$$+ \frac{2aa'(b)}{U'(b)} (\mu_3 - 2\mu_2^2 + \mu_2).$$

It now also follows from the central limit theorem that

$$\mathcal{L} (\sqrt{m}\ (\hat{s}(y) - s(y)) \to N\left(0, \sigma^2_{\hat{s}(y)}\right),$$

where $\sigma_{\hat{s}(y)}$ is too messy to justify writing down here. Observe that if $y, z > 0$ are such that

$$s(y) = \frac{(x_0+y)^b - x_b^b}{a} = z,$$

then

$$(x_0+y)^b = x_0^b + az$$

and

$$(x_0+y)^b \log(x_0+y) - x_0^b \log x_0$$

$$= az \log x_0 + \frac{x_0^b}{b} \left(1+ \frac{az}{x_0^b}\right) \log \left(1+ \frac{az}{x_0^b}\right) .$$

In particular

$$(x_0+Y)^b \log(x_0+Y) - x_0^b \log x_0$$

$$= aZ \log x_0 + \frac{x_0^b}{b} \left(1+ \frac{aZ}{x_0^b}\right) \log\left(1 + \frac{aZ}{x_0^b}\right) .$$

Since

$$\log(1+t) - 1 + \frac{t^2}{2} = t^2\, O(\min(t,1)), \qquad t>0,$$

it follows from the hypotheses of Theorem 5.1 that for k=0,1,2 (as $x_0 \to \infty$)

$$E\left[Z^k\left(\log\left(1+\frac{aZ}{x_0^b}\right) - \frac{aZ}{x_0^b} + \frac{1}{2}\left(\frac{aZ}{x_0^b}\right)^2\right]\right] = o\left(\left(\frac{a}{x_0^b}\right)^2\right).$$

Consequently

$$a'(b) = E((x_0+Y)^b \log(x_0+Y) - x_0^b \log x_0)$$

$$= a(\log x_0 + \frac{1}{b}) + \frac{a^2}{bx_0^b} + o\left(\frac{a^2}{bx_0^b}\right)$$

and

$$E[Z((x_0+Y)^b \log(x_0+Y) - x_0^b \log x_0)]$$

$$= a\mu_2(\log x_0 + \frac{1}{b}) + \frac{3a^2}{bx_0^b} + o\left(\frac{a^2}{bx_0^b}\right).$$

It now follows easily that

$$\frac{bx_0^b}{2a} U'(b) \to 1$$

and hence from the above formula for $\sigma_{\hat{b}}$ that (5.3) holds. It also follows from the above formula for $\sigma_{\hat{a}}^2$ that (5.4) holds.

As $m\to\infty$ for fixed x_0

$$\hat{s}(y(z)) - z$$

$$= \left[bz \log x_0 + \frac{x_0^b}{a}\left(1+\frac{az}{x_0^b}\right)\log\left(1+\frac{az}{x_0^b}\right)\right]\left(\frac{\hat{b}-b}{b}\right) - z\frac{\hat{a}-a}{a} + o_p\left(\frac{1}{m^{\frac{1}{2}}}\right) .$$

Let (W_1,W_2) have a bivariate normal distribution with $EW_1 = EW_2 = 0$,

$$\text{Var } W_1 = \mu_4 - 4\mu_2\mu_3 + 4\mu_2^3 - \mu_2^2,$$

$$\text{Var } W_2 = \mu_2 - 1,$$

and

$$\text{Cov}(W_1,W_2) = \mu_3 - 2\mu_2^2 + \mu_2.$$

Then

$$\sigma^2_{\hat{s}(y(z))} = \text{Var}(f(z)W_1 - zW_2),$$

where
$$f(z) = \frac{1}{U'(b)}\left(z \log x_0 + \frac{x_0^b}{ab}\left(1+\frac{az}{x_0^b}\right)\log\left(1+\frac{az}{x_0^b}\right) - \frac{a'(b)}{a}\, z\right).$$

Now (as $x_0\to\infty$)

$$f(z) \to \frac{z}{2}\left(\frac{z}{2} - 1\right).$$

It follows from the hypotheses of Theorem 5.1 that Var $W_1\to 4$, Var $W_2\to 4$ and $\text{cov}(W_1,W_2)\to 0$. Therefore (5.5) holds, which completes the proof of Theorem 5.1.

Suppose the Weibull tail model is exactly valid for the x_0 under consideration. Let f denote the density of Y and let

$$I = \begin{pmatrix} I_{bb} & I_{bc} \\ I_{bc} & I_{cc} \end{pmatrix}$$

denote the Fisher information matrix, defined by

$$I_{bb} = -\int \frac{\partial^2 \log f}{\partial b^2} \, fdy \ ,$$

$$I_{bc} = -\int \frac{\partial^2 \log f}{\partial b \partial c} \, fdy \ ,$$

and

$$I_{cc} = -\int \frac{\partial^2 \log f}{\partial c^2} \, fdy \ .$$

Set $z_0 = (x_0/c)^b$ and let $E_1(z_0)$ denote the exponential integral given by

$$E_1(z_0) = \int_{z_0}^{\infty} \frac{e^{-z}}{z} \, dz \ .$$

It follows by straightforward computations that

$$I_{bb} = \frac{1}{b^2} \left(1 + \int_{z_0}^{\infty} (z \log^2 z - z_0 \log^2 z_0) \, e^{-(z-z_0)} dz\right) ,$$

$$I_{bc} = -\frac{1}{c}(\log z_0 + 1 + e^{z_0} E_1(z_0)),$$

and

$$I_{cc} = \frac{b^2}{c^2} \ .$$

Let I^{-1} denote the inverse Fisher information matrix, so that

$$I^{-1} = \begin{pmatrix} I^{-1}_{bb} & I^{-1}_{bc} \\ I^{-1}_{bc} & I^{-1}_{cc} \end{pmatrix} = \frac{1}{I_{bb}I_{cc} - (I_{bc})^2} \begin{pmatrix} I_{cc} & -I_{bc} \\ -I_{bc} & I_{bb} \end{pmatrix} .$$

Then

$$I_{bb}I_{cc} - (I_{bc})^2 = \frac{D(z_0)}{c^2} ,$$

where

$$D(z_0) = 1 + \int_{z_0}^{\infty} (z \log^2 z - z_0 \log^2 z_0)\, e^{-(z-z_0)}\, dz$$

$$- (\log z_0 + 1 + e^{z_0} E_1(z_0))^2 .$$

Let $\bar{b}$, $\bar{c}$ and $\bar{s}(y)$ be consistent maximum likelihood estimators of b, c and s(y) respectively based on the random sample $Y_1, \ldots, Y_m$. The asymptotic joint distribution of $\sqrt{m}(\bar{b}-b)$ and $\sqrt{m}(\bar{c}-c)$ is a bivariate normal distribution with mean 0 and covariance matrix I^{-1}.

Observe that (as $x_0 \to \infty$ or, equivalently, as $z_0 \to \infty$)

$$e^{z_0} E_1(z_0) = \frac{1}{z_0} - \frac{1}{z_0^2} + 0\left(\frac{1}{z_0^3}\right)$$

and hence

$$I_{bc} = -\frac{1}{c}\left[\log z_0 + 1 + \frac{1}{z_0} - \frac{1}{z_0^2} + 0\left(\frac{1}{z_0^3}\right)\right].$$

By straightforward calculations

$$I_{bb} = \frac{1}{b^2}\left[(\log z_0 + 1)^2 + (\log z_0 + 1)\left(\frac{2}{z_0} - \frac{2}{z_0^2}\right) + \frac{2}{z_0^2} + 0\left(\frac{\log z_0}{z_0^3}\right)\right].$$

Consequently

$$(cz_0)^2(I_{bb}I_{cc} - (I_{bc})^2) \to 1 \quad .$$

Since

$$\sigma_b^2 = I^{-1}_{bb} = \frac{I_{cc}}{I_{bb}I_{cc}-(I_{bc})^2} = \frac{b^2/c^2}{I_{bb}I_{cc}-(I_{bc})^2} \quad ,$$

(5.6) is valid.

As $m \to \infty$ for fixed $\mathring{x}_0$

$$\mathcal{L}(\sqrt{m}\ (\bar{s}(y) - s(y))) \to N(0, \sigma^2_{\bar{s}(y)}) \ ,$$

where

$$\sigma^2_{\bar{s}(y)} = \left[\left(\frac{x_0+y}{c}\right)^b \log\frac{x_0+y}{c} - \left(\frac{x_0}{c}\right)^b \log\frac{x_0}{c}\right]^2 I^{-1}_{bb}$$

$$- 2\frac{b}{c}\left[\left(\frac{x_0+y}{c}\right)^b \log\frac{x_0+y}{c} - \left(\frac{x_0}{c}\right)^b \log\frac{x_0}{c}\right] s(y) I^{-1}_{bc}$$

$$+ \frac{b^2}{c^2} s^2(y) I^{-1}_{cc} \quad .$$

Let $y=y(z)$ be defined in terms of $z>0$ so that

$$s(y) = \left(\frac{x_0+y}{c}\right)^b - \left(\frac{x_0}{c}\right)^b = z.$$

As x_0, $z_0 \to \infty$

$$b\left[\left(\frac{x_0+y}{c}\right)^b \log \frac{x_0+y}{c} - \left(\frac{x_0}{c}\right)^b \log \frac{x_0}{c}\right]$$

$$= z \log z_0 + (z_0+z) \log\left(1+ \frac{z}{z_0}\right)$$

$$= z\left(\log z_0 + 1 + \frac{z}{2z_0} - \frac{z^2}{6z_0^2}\right) + 0\left(\frac{1}{z_0^3}\right) .$$

Equation (5.7) now follows by a straightforward calculation.

By (5.3) and 5.6) as $x_0 \to \infty$

$$\frac{\sigma_{\hat{b}}}{\sigma_{\overline{b}}} \to 1 \quad .$$

In this sense $\hat{b}$ is an asymptotically efficient estimator of b as first $m \to \infty$ and then $x_0 \to \infty$. Similarly it follows from (5.5) and (5.7) that $\hat{s}(y(z))$ is an asymptotically efficient estimator of $s(y(z))$ as first $m \to \infty$ and then $x_0 \to \infty$.

6. ASYMPTOTIC BEHAVIOR OF EXPONENTIAL TAIL ESTIMATORS

Recall that X is a positive random variable having distribution function F and that $\ell(x) = -\log P(X>x) = -\log (1-F(x)) = -\log S(x)$, $x>0$. It is supposed throughout this section that ℓ is twice differentiable and $\ell'>0$ on $(0,\infty)$. Recall also that $r = \ell''/(\ell')^2$ on $(0,\infty)$. Results concerning the asymptotic behavior of exponential tail estimators of a quantile $x_{p/n}$ are stated in Section 6.1 and proven in Section 6.2.

6.1 DISCUSSION OF RESULTS

Suppose that

$$\lim_{x\to\infty} r(x) = 0 \quad . \tag{6.1}$$

Then

$$\lim_{x\to\infty} x\ell'(x) = \infty \quad , \tag{6.2}$$

so

$$\lim_{x\to\infty} (\ell(x+y/\ell'(x)) - \ell(x) = y, \quad -\infty<y<\infty \quad , \tag{6.3}$$

and hence

$$\lim_{x\to\infty} \frac{S(x+y/\ell'(x))}{S(x)} = e^{-y}, \quad -\infty<y<\infty \quad . \tag{6.4}$$

As pointed out in Section 2.3 in connection with the discussion of heaviness, (6.1) holds for Weibull and lognormal distributions, but the convergence is very slow. Similarly there is slow convergence in (6.4) or, equivalently, in

$$\lim_{p\to\infty} p^{-1} S(x_p+y/\ell'(x_p)) = e^{-y}, \quad -\infty<y<\infty \quad .$$

Recall that $M_n = X_1 = \max(X_1,\ldots,X_n)$. It follows from (6.4) that

$$\lim_{n\to\infty} F^n(x_{1/n}+y/\ell'(x_{1/n}))$$

$$= \lim_{n\to\infty} P(\ell'(x_{1/n})(M_n-X_{1/n}) \le y) = e^{-e^{-y}}, \quad -\infty<y<\infty \quad . \tag{6.5}$$

Proofs that (6.2) - (6.5) follow from (6.1) will be given in Section 6.2. The classical result that (6.1) implies (6.5) is due to von Mises [10].

Given a positive integer m, let G_m denote the distribution function given by

$$G_m(y) = \sum_{0}^{m-1} \frac{1}{k!} e^{-ky} e^{-e^{-y}}, \quad -\infty<y<\infty \quad ;$$

its density is given by

$$\frac{d}{dy} G_m(y) = \frac{e^{-my}}{(m-1)!} e^{-e^{-y}}, \quad -\infty<y<\infty \quad .$$

(If Y has distribution function G_m, then e^{-Y} has a gamma distribution with parameters m and 1.) The moments of G_m are all finite. Let μ_m and σ_m^2 denote respectively the mean and variance of G_m. Then

$$\mu_1 = \int_{-\infty}^{\infty} y\, e^{-y}\, e^{-e^{-y}}\, dy = -\int_0^{\infty} \log z\, e^{-z}\, dz = \gamma \quad ,$$

where $\gamma \approx .5772$ is Euler's constant. Also

$$\int_{-\infty}^{\infty} y^2\, e^{-y}\, e^{-e^{-y}}\, dy = \int_0^{\infty} \log^2 z\, e^{-z}\, dz = \frac{\pi^2}{6} + \gamma^2 \quad ,$$

so $\sigma_1^2 = \pi^2/\delta$. It is easily seen that

$$\int_{-\infty}^{z} \frac{1}{m!}\, e^{-(m+1)y}\, e^{-e^{-y}}\, me^{-m(z-y)}\, dy = \frac{1}{(m-1)!}\, e^{-mz}\, e^{-e^{-z}} \quad ,$$

so that $G^{(m)}$ is the convolution of $G^{(m+1)}$ with an exponential distribution having mean $1/m$ and variance $1/m^2$. Consequently $\mu_m = \mu_{m+1} + 1/m$ and $\sigma_m^2 = \sigma_{m+1}^2 + 1/m^2$ and therefore

$$\mu_m = \gamma - \sum_1^{m-1} \frac{1}{k} \text{ and } \sigma_m^2 = \frac{\pi^2}{6} - \sum_1^{m-1} \frac{1}{k^2} = \sum_m^{\infty} \frac{1}{k^2} \quad .$$

In particular the variance of G_m approaches zero as $m \to \infty$.

Recall that X_m is the m^{th} largest value among $X_1, \ldots, X_n$ (the dependence of X_m on n is suppressed). The next result, which strengthens and generalizes (6.5), will be proven in Section 6.2.

Theorem 6.1. Suppose (6.1) holds. Then as $n\to\infty$ the distribution function of $\ell'(x_{1/n})(X_m - x_{1/n})$ converges to G_m and the moments of this random variable converge to the corresponding moments of G_m.

Given a positive integer m and a positive number p with $0<p\leq m+1$, let $\hat{x}^{(m)}_{p/n}$ denote the exponential tail estimator of the quantile $x_{p/n}$ based on a random sample of size n as given by (2.3); so that

$$\hat{x}^{(m)}_{p/n} = X_{m+1} + \hat{a}^{(m)} \log \frac{m+1}{p} \quad , \tag{6.6}$$

where

$$\hat{a}^{(m)} = \frac{1}{m} \sum_{1}^{m} (X_i - X_{m+1}) \quad . \tag{6.7}$$

The asymptotic behavior of $\hat{x}^{(m)}_{p/n}$ as $n\to\infty$ for fixed m is described in the next result, which will be proven in Section 6.2.

Theorem 6.2. Suppose (6.1) holds and $0<p\leq m+1$. Then as $n\to\infty$ the distribution function and moments of $\ell'(x_{1/n})(\hat{x}^{(m)}_{p/n} - x_{p/n})$ converge to those of a random variable $Z^{(m)}_p$ which has finite moments of all orders and positive variance; moreover $E(Z^{(m)}_p)^2 \to 0$ as $m\to\infty$.

The following easily obtained consequence of Theorem 6.1 and Theorem 6.2 will also be proven in Section 6.2.

Theorem 6.3. Suppose (6.1) holds and $p>0$. If $\{m_n\}$ is a sequence of positive integers which tends to infinity sufficiently slowly as $n\to\infty$, then

$$\lim_{n\to\infty} \frac{E\left(\hat{x}_{p/n}^{(m_n)} - x_{p/n}\right)^2}{\mathrm{Var}(\hat{x}_{p/n}^{(m)})} = 0, \quad 0<p\leq m+1 \quad ,$$

and

$$\lim_{n\to\infty} \frac{E\left(\hat{x}_{p/n}^{(m_n)} - x_{p/n}\right)^2}{\mathrm{Var}\ X_m} = 0 \quad ;$$

in particular

$$\lim_{n\to\infty} \frac{E\left(\hat{x}_{p/n}^{(m_n)} - x_{p/n}\right)^2}{\mathrm{Var}\ M_n} = 0 \quad .$$

Theorem 6.1 and Theorem 6.3 together yield the following result, which shows that under appropriate conditions $\hat{x}_{p/n}^{(m_n)}$ is a consistent estimator of $x_{p/n}$.

Theorem 6.4. Suppose that $p>0$, that (6.1) holds and that

$$\lim_{x\to\infty} \ell'(x) = \infty \quad . \tag{6.8}$$

If $\{m_n\}$ is a sequence of positive integers which tends to infinity sufficiently slowly as n , then

$$\lim_{n\to\infty} E\left(\hat{x}_{p/n}^{(m_n)} - x_{p/n}\right)^2 = 0 \quad .$$

Note that (6.8) holds for such light-tailed distributions as Weibull distributions with shape parameter $b>1$. But it fails for medium-tailed

and heavy-tailed distributions in which $r(x) \le 0$ for x sufficiently large; in particular, it fails for exponential distributions, for Weibull distributions with shape parameter $b<1$, and for lognormal distributions.

6.2 PROOFS

Suppose from now on that (6.1) holds. It will first be shown that (6.2) holds. Suppose otherwise. Then there is a positive number M and there is a sequence $\{x_j\}$ of positive numbers tending to infinity such that $x_j \ell'(x_j) \le M$ for $j \ge 1$. By (6.1) there is a positive number x_0 such that

$$\frac{d}{dx}\frac{1}{\ell'(x)} = -r(x) \le \frac{1}{2M}\ , \ x \ge x_0 \quad .$$

Choose j such that $x_j > x_o$. By the mean value theorem

$$\frac{1}{\ell'(x_j)} - \frac{1}{\ell'(x)} \le \frac{x_j - x}{2M} \le \frac{x_j}{2M}\ , \ x_o \le x \le x_j \quad ,$$

and hence

$$\ell'(x) \le \frac{2M}{x_j}\ , \ x_o \le x \le x_j \quad .$$

Consequently $\ell(x_j) \le \ell(x_o) + 2M$ for $j \ge 1$, which contradicts the fact that $\ell(x) = -\log P(X>x) \to \infty$ as $x \to \infty$. Therefore (6.2) holds as desired.

Next (6.3) will be verified. By the mean-value theorem it suffices to show that for $M>0$

$$\lim_{x\to\infty} \frac{\ell'(x+y/\ell'(x))}{\ell'(x)} = 1 \text{ uniformly for } |y| \leq M \quad . \tag{6.9}$$

Choose $\varepsilon > 0$. By (6.1), (6.2) and the mean-value theorem there is an $x_0 > 0$ such that if $x \geq x_0$ and $|y| \leq M$, then

$$\left| \frac{1}{\ell'(x+y/\ell'(x))} - \frac{1}{\ell'(x)} \right| \leq \frac{\varepsilon|y|}{\ell'(x)} \leq \frac{\varepsilon M}{\ell'(x)}$$

and hence

$$\left| \frac{\ell'(x)}{\ell'(x+y/\ell'(x))} - 1 \right| \leq \varepsilon M \quad .$$

Therefore (6.9) is valid and hence so is (6.3). Now

$$\frac{S(x+y/\ell'(x))}{S(x)} = \exp[-(\ell(x+y/\ell'(x))-\ell(x)] \quad ,$$

so (6.4) follows immediately from (6.3). Also

$$F^n(x_{1/n}+y/\ell'(x_{1/n})) = \left(1 - \frac{1}{n}\,\frac{S(x_{1/n}+y/\ell'(x_{1/n}))}{S(x_{1/n})}\right)^n \quad ,$$

so (6.5) follows from (6.4). This completes the proof that (6.2) - (6.5) follow from (6.1).

In preparation for the proof of Theorem 6.1, two further elementary consequences of (6.1) will now be obtained. Let $c>1$. Then

$$\lim_{p\to o} \frac{x_{cp}-x_{c^2p}}{x_p-x_{cp}} = 1 \tag{6.10}$$

and

$$\lim_{p\to o} \ell'(x_p)(x_p-x_{cp}) = \log c \quad . \tag{6.11}$$

To prove these results, choose ε with $0<\varepsilon<1$. By (6.1) there is an $x_o>0$ such that

$$\left|\frac{1}{\ell'(x+y)} - \frac{1}{\ell'(x)}\right| \leq \varepsilon^2 y \text{ for } x \geq x_o \text{ and } y \geq 0 \quad .$$

In particular

$$\left|\frac{1}{\ell'(x+y)} - \frac{1}{\ell'(x)}\right| \leq \frac{\varepsilon}{\ell'(x)} \text{ for } x \geq x_o \text{ and } 0 \leq y \leq \frac{1}{\varepsilon\ell'(x)} \quad .$$

Consequently

$$\frac{\ell'(x)}{1+\varepsilon} \leq \ell'(x+y) \leq \frac{\ell'(x)}{1-\varepsilon} \text{ for } x \geq x_o \text{ and } 0 \leq y \leq \frac{1}{\varepsilon\ell'(x)} \quad .$$

Therefore

$$\frac{S(x)}{S(x+1/\varepsilon\ell'(x))} = \exp(\ell(x+1/\varepsilon\ell'(x)) - \ell(x)) \geq e^{1/\varepsilon(1+\varepsilon)} \quad .$$

Choose $c>1$ and let ε be small enough that $e^{1/\varepsilon(1+\varepsilon)} \geq c^2$. Choose $p_o>0$ such that $x_{c^2p_o} \geq x_o$. Then for $p \geq p_o$

$$\frac{\ell'(x_{c^2p})}{1+\varepsilon} \leq \ell'(z) \leq \frac{\ell'(x_{c^2p})}{1-\varepsilon} \text{ for } x_{c^2p} \leq z \leq x_p \quad .$$

Also

$$c = \frac{S(x_{c^2p})}{S(x_{cp})} = e^{\ell(x_{cp}) - \ell(x_{c^2p})} \quad .$$

Thus $\log c = \ell'(z)(x_{cp}-x_{c^2p})$ for some z between x_{c^2p} and x_{cp} and hence

$$\frac{(1-\varepsilon)\log c}{\ell'(x_{c^2p})} \leq x_{cp} - x_{c^2p} \leq \frac{(1+\varepsilon)\log c}{\ell'(x_{c^2p})} \text{ for } p \geq p_o \quad .$$

Similarly

$$\frac{(1-\varepsilon)\log c}{\ell'(x_{c^2p})} \leq x_p - x_{cp} \leq \frac{(1+\varepsilon)\log c}{\ell'(x_{c^2p})} \text{ for } p \geq p_o$$

and consequently

$$\frac{1-\varepsilon}{1+\varepsilon} \leq \frac{x_{cp}-x_{c^2p}}{x_p-x_{cp}} \leq \frac{1+\varepsilon}{1-\varepsilon} \text{ for } p \geq p_o \quad .$$

Since ε can be made arbitrarily small, (6.10) is valid. It is now clear that (6.11) is also valid.

Theorem 6.1 will now be verified. First the well known result that the distribution function of $\ell'(x_{1/n})(X_m-x_{1/n})$ converges to G_m will be obtained. To this end, observe that

$$P(X_m \leq x) = \sum_{0}^{m-1} \binom{n}{k}(1-F(x))^k F^{n-k}(x)$$

and hence

$$P(\ell'(x_{1/n})(X_m-x_{1/n}) \leq y)$$

$$= \sum_{0}^{m-1} n^{-k}\binom{n}{k}\left(\frac{S(x_{1/n}+y/\ell'(x_{1/n}))}{S(x_{1/n})}\right)^k F^{n-k}(x_{1/n}+y/\ell'(x_{1/n})) \quad .$$

Since $\lim_n n^{-k}\binom{n}{k} = 1/k!$, it now follows from (6.4) and (6.5) that

$$\lim_{n\to\infty} P(\ell'(x_{1/n})(X_m-x_{1/n})\leq y) = \sum_{0}^{m-1} \frac{1}{k!} e^{-ky} e^{-e^{-y}} = G_m(y), \quad -\infty<y<\infty \quad ,$$

as desired.

It remains to prove that the (positive integer) moments of $\ell'(x_{1/n})(X_m - x_{1/n})$ converge to those of G_m. For m=1 this result has been obtained in a more general context by Pickands [11]. It is possible to reduce the result for $m \geq 1$ to that for m=1. But since Pickand's proof is rather complicated, a self-contained proof for $m \geq 1$ will be given.

Since convergence in distribution has already been established, it suffices to show that the moments of $\ell'(x_{1/n})(X_m - x_{1/n})$ are bounded in n. For this it suffices to show that for each positive integer j

$$T_{n1} = (\ell'(x_{1/n}))^j \int_0^{x_{1/n}} y^{j-1} P(X_m \leq x_{1/n} - y)dy$$

and

$$T_{n2} = (\ell'(x_{1/n}))^j \int_0^{\infty} y^{j-1} P(X_m > x_{1/n} + y)dy$$

are bounded in n. To verify that T_{n1} is bounded in n, it is enough to show that for $0 \leq k \leq m-1$

$$T_{n3} = (\ell'(x_{1/n}))^j \int_0^{x_{1/n}} (x_{1/n} - y)^{j-1} (nS(y))^k (1-S(y))^{n-k} dy$$

is bounded in n.

Choose δ with $0 < \delta \leq 1/2$. Choose i(n) such that (for n sufficiently large)

$$\delta^2 n \leq 2^{i(n)} \leq \delta n \quad .$$

Then

$$T_{n3} = (\ell'(x_{1/n}))^j \sum_{1}^{i(n)} \int_{x_{2^i/n}}^{x_{2^{i-1}/n}} (x_{1/n}-y)^{j-1}(nS(y))^k(1-S(y))^{n-k}dy$$

$$+ (\ell'(x_{1/n}))^j \int_{0}^{x_{2^{i(n)}/n}} (x_{1/n}-y)^{j-1}(nS(y))^k(1-S(y))^{n-k}dy \quad .$$

Observe that if $x_{2^i/n} \leq y \leq x_{2^{i-1}/n}$, then

$$(x_{1/n}-y)^{j-1}(nS(y))^k(1-S(y))^{n-k}$$

$$\leq (x_{1/n}-x_{2^i/n})^{j-1}(nS(x_{2^i/n}))^k(1-S(x_{2^{i-1}/n}))^{n-k}$$

$$= (x_{1/n}-x_{2^i/n})^j 2^{ik}\left(1-\frac{2^{i-1}}{n}\right)^n \left(1-\frac{2^{i-1}}{n}\right)^{-k}$$

$$\leq (x_{1/n}-x_{2^i/n})^j 2^{ik} \cdot e^{-2^{i-1}} 2^{-k} \quad .$$

Observe also that if $0 \leq y \leq x_{2^i/n}$, then

$$(x_{1/n}-y)^{j-1}(nS(y))^k(1-S(y))^{n-k}$$

$$\leq (x_{1/n})^{j-1} n^k\left(1-\frac{2^{i(n)}}{n}\right)^{n-k}$$

$$\leq (x_{1/n})^{j-1} n^k e^{-2^{i(n)}} 2^{-k}$$

$$\leq (x_{1/n})^{j-1} n^k e^{-\delta^2 n} 2^{-k} \quad .$$

Thus to prove that T_{n3} is bounded in n, it suffices to show that

$$T_{n4} = (\ell'(x_{1/n}))^j \sum_{1}^{i(n)} (x_{2^{i-1}/n}-x_{2^i/n})(x_{1/n}-x_{2^i/n})^{j-1}\, 2^{ki}\, e^{-2^{i-1}}$$

and

$$T_{n5} = (\ell'(x_{1/n}))^j (x_{1/n})^{j-1}\, n^k\, e^{-\delta^2 n}$$

are bounded in n.

Choose $\varepsilon > 0$. By (6.10) and (6.11) if δ is made sufficiently small, then (for n sufficiently large)

$$\frac{x_{2^i/n}-x_{2^{i+1}/n}}{x_{2^{i-1}/n}-x_{2^i/n}} \leq 1 + \varepsilon, \quad 1\leq i\leq i(n) \quad ,$$

and

$$\ell'(x_{1/n})(x_{1/n}-x_{2/n}) \leq 1 \quad .$$

Consequently

$$x_{2^{i-1}/n}-x_{2^i/n} \leq \frac{(1+\varepsilon)^{i-1}}{\ell'(x_{1/n})}\, , \quad 1\leq i\leq i(n) \quad ,$$

hence

$$x_{1/n}-x_{2^i/n} \leq \frac{1}{\ell'(x_{1/n})} \frac{(1+\varepsilon)^i}{\varepsilon}\, , \quad 1\leq i\leq i(n) \quad ,$$

and therefore

$$T_{n4} \le \sum_{1}^{i/n} (1+\varepsilon)^{2i}\ 2^{ki}\ e^{-2^{i-1}} \qquad ,$$

which is bounded in n. Similarly it follows from (6.10) and (6.11) that

$$\ell'(x_{1/n}) = O((1+\varepsilon)^{\log_2 n}) = O(n^{\log_2(1+\varepsilon)}$$

and hence that

$$T_{n5} = O(n^{\log_2(1+\varepsilon)+k}\ e^{-\delta^2 n}) \qquad .$$

Thus T_{n5} is also bounded in n. This completes the proof that T_{n1} is bounded in n.

To prove that T_{n2} is bounded in n, it suffices to verify the boundedness in n of

$$T_{n6} = (\ell'(x_{1/n}))^j \int_0^\infty y^{j-1}\ (1-(1-S(x_{1/n}+y))^n)dy$$

$$= (\ell'(x_{1/n}))^j \sum_1^\infty \int_{x_{2^{1-i}/n}}^{x_{2^{-i}/n}} (y-x_{1/n})^{j-1}(1-S(y))$$

$$\le (\ell'(x_{1/n}))^j \sum_1^\infty (x_{2^{-i}/n}-x_{2^{1-i}/n})(x_{2^{-i}/n}-x_{1/n})^j\ (1-(1-2^{1-i}/n)^n)dy$$

$$= O(\ell'(x_{1/n}))^j \sum_1^\infty (x_{2^{-i}/n}-x_{2^{1-i}/n})(x_{2^{-i}/n}-x_{1/n})^j\ 2^{-i}) \qquad .$$

By an argument similar to that used in verifying the boundedness in n of T_{n4}, it now follows that T_{n6} is bounded in n. This completes the proof of Theorem 6.1.

In preparation for the proof of Theorem 6.2, another elementary consequence of (6.1) will be obtained:

$$\lim_{n\to\infty} \ell'(x_{1/n})(F^{-1}(1-\frac{z}{n})-x_{1/n}) = \log \frac{1}{z},\ z>0 \quad . \tag{6.12}$$

Choose $\varepsilon > 0$. By (6.4)

$$\lim_{n\to\infty} n\left(1-F\left(x_{1/n} + \frac{\log(1/z)+\varepsilon}{\ell'(x_{1/n})}\right)\right) = e^{-\varepsilon}z$$

and

$$\lim_{n\to\infty} n\left(1-F\left(x_{1/n} + \frac{\log(1/z)-\varepsilon}{\ell'(x_{1/n})}\right)\right) = e^{\varepsilon}z \quad ,$$

so for n sufficiently large

$$\log \frac{1}{z} = \varepsilon \leq \ell'(x_{1/n})(F^{-1}(1-\frac{z}{n})-x_{1/n}) \leq \log \frac{1}{z} + \varepsilon \quad .$$

Since ε can be made arbitrarily small, (6.12) holds as desired.

Theorem 6.2 will now be verified. Let $W_1,\ldots,W_m$, $Z_{m+1}^{(n)}$ be independent variables such that W_i has an exponential distribution with mean 1 for each i and

$$P(Z_{m+1}^{(n)} \leq x) = \sum_{0}^{m-1} \binom{n}{k} e^{-kx}(1-e^{-x})^{n-k}, \ x \geq 0 \qquad ;$$

let $W_1,\ldots,W_m$, Z_{m+1} be independent and suppose that Z_{m+1} has distribution function G_{m+1}. Then $Z_{m+1}^{(n)}$ - log n converges in distribution to Z_{m+1} as $n\to\infty$. For $1\leq i\leq m$ set

$$Z_i^{(n)} = Z_{m+1}^{(n)} + \sum_{i}^{m} \frac{W_j}{j}$$

and

$$Z_i = Z_{m+1} + \sum_{i}^{m} \frac{W_j}{j} \quad .$$

Then $Z_i^{(n)},\ldots,Z_{m+1}^{(n)}$ are well known to be jointly distributed as the m+1 largest values in a random sample of size n from an exponential distribution with mean one; and as $n\to\infty$ the joint distribution of $Z_1^{(n)}$ - log n,..., $Z_{m+1}^{(n)}$ - log n converges to that of $Z_1,\ldots,Z_{m+1}$. By the first result in the last sentence

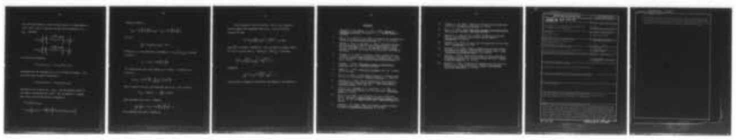
$$1-e^{-Z_1^{(n)}},\ldots,1-e^{-Z_{m+1}^{(n)}}$$

are jointly distributed as the m+1 largest values in a random sample of size n from the uniform distribution on [0,1]. Consequently

$$F^{-1}\left(1 - \frac{e^{-(Z_1^{(n)}-\log n)}}{n}\right),\ldots,F^{-1}\left(1 - \frac{e^{-(Z_{m+1}^{(n)}-\log n}}{n}\right)$$

AD-A080 446 TECHNOLOGY SERVICE CORP SANTA MONICA CALIF F/G 12/1
NEW METHODS FOR ESTIMATING TAIL PROBABILITIES AND EXTREME VALUE--ETC(U)
DEC 79 L BREIMAN, C J STONE, J D GINS F49620-79-C-0171
UNCLASSIFIED TSC-PD-A226-1 AFOSR-TR-80-0019 NL

2 OF 2
AD
A080446



END
DATE
FILMED
3 - 80
DDC

are jointly distributed as the m+1 largest values in a random sample of size n from F; that is, they have the same joint distribution as $X_1, \ldots, X_{m+1}$. Therefore

$$\ell'(x_{1/n})\left(F^{-1}\left(1 - \frac{e^{-(Z_1^{(n)} - \log n)}}{n}\right) - x_{1/n}\right), \ldots,$$

$$\ell'(x_{1/n})\left(F^{-1}\left(1 - \frac{e^{-(Z_{m+1}^{(n)} - \log n)}}{n}\right) - x_{1/n}\right)$$

are jointly distributed as

$$\ell'(x_{1/n})(X_1 - x_{1/n}), \ldots, \ell'(x_{1/n})(X_{m+1} - x_{1/n}) \quad .$$

By monotonicity the convergence in (6.12) is uniform on compacts. This now implies that the joint distribution of

$$\ell'(x_{1/n})(X_1 - x_{1/n}), \ldots, \ell'(x_{1/n})(X_{m+1} - x_{1/n})$$

converges as $n \to \infty$ to that of $Z_1, \ldots, Z_{m+1}$. (For more general results to this effect, see Weissman [17], [18].) This and Theorem 6.1 together imply that as $n \to \infty$ the distribution and moments of

$$\ell'(x_{1/n})(\hat{x}_{p/n} - x_{1/n})$$

$$= \ell'(x_{1/n})\left[X_{m+1} - x_{1/n} + \log\frac{m+1}{p}\,\frac{1}{m}\sum_{1}^{m}((X_i - x_{1/n}) - (X_{m+1} - x_{1/n}))\right]$$

converge to those of

$$Z_{m+1} + \log \frac{m+1}{p} \frac{1}{m} \sum_{1}^{m} (Z_i - Z_{m+1}) = Z_{m+1} + \log \frac{m+1}{p} \frac{1}{m} \sum_{1}^{m} W_i \quad .$$

By (6.11)

$$\lim_{n\to\infty} \ell'(x_{1/n})(x_{1/n} - x_{p/n}) = \log p \quad .$$

Therefore as $n\to\infty$ the distribution and moments of $\ell'(x_{1/n})(\hat{x}_{p/n} - x_{p/n})$ converge to those of

$$Z_{m+1} + \log p + \log \frac{m+1}{p} \frac{1}{m} \sum_{1}^{m} W_i \quad .$$

This random variable has finite moments of all orders. In particular its variance is

$$\mathrm{Var}\, Z_{m+1} + \frac{1}{m} \log^2 \frac{m+1}{p} = \sum_{m+1}^{\infty} \frac{1}{k^2} + \frac{1}{m} \log^2 \frac{m+1}{p} \quad ,$$

which is positive for each m and approaches zero as $m\to\infty$. Also its mean is

$$E\, Z_{m+1} + \log(m+1) = \gamma - \sum_{1}^{m} \frac{1}{k} + \log(m+1) \quad ,$$

which approaches zero as $m\to\infty$. Therefore

$$\lim_{n\to\infty} E\left(Z_{m+1} + \log p + \log \frac{m+1}{p} \frac{1}{m} \sum_{1}^{m} W_i\right)^2 = 0 \quad ,$$

which completes the proof of Theorem 6.2.

Finally Theorem 6.3 will be verified. Let $\{\varepsilon_m\}$ be a sequence of positive numbers which approaches zero as $m\to\infty$. Let n_m be positive integers such that

$$E(\ell'(x_{1/n})(\hat{x}_{p/n}^{(m)}-x_{p/n}))^2 \leq E(Z_p^{(m)})^2 + \varepsilon_m, \ n \geq n_m \quad ,$$

where $Z_p^{(m)}$ is defined in Theorem 6.2. Let m_n be positive integers tending to ∞ as $n\to\infty$ and such that $m_n < \inf[m; n_m > n]$. Then $n_{m_n} \leq n$ and hence

$$E\left(\ell'(x_{1/n})\left(\hat{x}_{p/n}^{(m_n)} - x_{p/n}\right)\right)^2 \leq E\left(Z_p^{(m_n)}\right)^2 + \varepsilon_n \quad .$$

Consequently

$$\lim_{n\to\infty} (\ell'(x_{1/n}))^2 \, E\left(\hat{x}_{p/n}^{(m_n)} - x_{p/n}\right)^2 = 0 \quad .$$

The conclusion of Theorem 6.3 now follows from Theorem 6.1 and Theorem 6.2.

REFERENCES

1. Abramowitz, M. and Stegun, I. A. (eds.) (1964), Handbook of Mathematical Functions, Washington, D. C. : National Bureau of Standards.

2. Box, G. E. P. and Cox, D. R. (1964), "An Analysis of Transformations," Journal of the Royal Statistical Society, Ser. B, 26, 211-252.

3. Breiman, L., Gins, J. and Stone, C. (1978), "Statistical Analysis and Interpretation of Peak Air Pollution Measurements," TSC-PD-A190-10, Final Report on U. S. Environmental Protection Agency, Contract 68-02-2857, Technology Service Corporation, Santa Monica, California.

4. Filimonov, V. A. (1978), "Estimation of Small Probabilities with the Help of a Generalized Distribution of Extreme Values," (translated from Russian), Radio Engineering and Electronic Physics, 23, No. 2, 151-153.

5. Galambos, J. (1978), The Asymptotic Theory of Extreme Order Statistics, New York: John Wiley and Sons.

6. Gumbel, E. J. (1978), Statistics of Extremes, New York: Columbia University Press.

7. Hill, B. M. (1975), "A Simple General Approach to Inference About the Tail of a Distribution," Annals of Statistics, 3, 1163-1174.

8. Jeruchim, M. C. (1976), "On the Estimation of Error Probability Using Generalized Extreme-Value Theory," IEEE Transactions on Information Theory, IT-22, 108-110.

9. Lewis, P. A. W., Goodman, A. S., and Miller, J. M. (1969), "A Pseudo-Random Number Generator for the System 1360," IBM Systems Journal, 8, 136-146.

10. Mises, R. von (1936), "La Distribution de la Plus Grande de n Valeur," Reprinted in Selected Papers of Richard von Mises (1954), Vol. II, 271-294, Providence, Rhode Island: American Mathematical Society.

11. Pickands, J. III (1968), "Moment Convergence of Sample Extremes," The Annals of Mathematical Statistics, 39, 881-889.

12. Pickands, J. III (1975), "Statistical Inference Using Extreme Order Statistics," The Annals of Statistics, 3, 119-131.

13. Rao, C. R. (1973), Linear Statistical Inference and its Applications, Second Edition, New York: John Wiley and Sons.

14. Shapiro, S. S., and Wilk, M. B. (1972), "An Analysis of Variance Test for the Exponential Distribution (Complete Samples)," Technometrics, 14, 355-370.

15. Stephens, M. A. (1978), "On the W Test for Exponentiality with Origin Known," Technometrics, 20, 33-35.

16. Weinstein, S. B. (1973), "Theory and Application of Some Classical and Generalized Asymptotic Distributions of Extreme Values," IEEE Transactions on Information Theory, IT-19, 148-154.

17. Weissman, I. (1975), "Multivariate Extremal Processes Generated by Independent Non-Identically Distributed Random Variables," Journal of Applied Probability, 12, 477-487.

18. Weissman, I. (1978), "Estimation of Parameters and Large Quantiles Based on the k Largest Observations," Journal of the American Statistical Association, 73, 812-815.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 80-0019	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) NEW METHODS FOR ESTIMATING TAIL PROBABILITIES AND EXTREME VALUE DISTRIBUTIONS		5. TYPE OF REPORT & PERIOD COVERED Final
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Leo Breiman, Charles J. Stone, John Gins		8. CONTRACT OR GRANT NUMBER(s) F49620-79-C-0171
9. PERFORMING ORGANIZATION NAME AND ADDRESS Technology Service Corporation, 2811 Wilshire Blvd., Santa Monica, CA 94003		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM, Bolling AFB, Washinton, DC 20332		12. REPORT DATE 14 December 1979
		13. NUMBER OF PAGES 101
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

tail probability, survival probability, extreme value distribution, quantile, exponential tail estimator, transformation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This research has focused on the problem of estimating probabilities in the upper tail of an underlying distribution and the corresponding quantiles based on a random sample from the distribution. Two estimation procedures, exponential tail and transformed exponential tail, were defined and their bias and variance properties were thoroughly studied both analytically and by means of an extensive Monte Carlo experiment. The experiment involved several forms of each of the two procedures; twenty underlying distributions were simulated, including a variety of Weibull and lognormal distributions; foursample sizes were

DD FORM 1 JAN 73 1473

Unclassified

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

20. Abstract cont.

considered--100, 200, 400 and 800. Careful study of the analytic and Monte Carlo results showed that exponential tail and transformed exponential tail procedures worked quite well, but indicated a potential for substantial further improvement by properly combining them.

UNCLASSIFIED